UNITED STATES DEPARTMENT OF COMMERCE
Jesse H. Jones, Secretary
NATIONAL BUREAU OF STANDARDS
Lyman J. Briggs, Director

✦

Journal of Research

of the

National Bureau of Standards

✦

Volume 26, Number 2

February 1941

[Published with approval of the Director of the Budget]

UNITED STATES
GOVERNMENT PRINTING OFFICE
WASHINGTON : 1941

For sale by the Superintendent of Documents, Washington, D. C. - - - - - Price 30 cents
$3.50 per year on subscription

The prices of the separate Research Papers appearing in this JOURNAL are given on the front cover page. If 100 or more copies of any separate are ordered, a discount of 25 percent is allowed. Those who desire copies of separate Research Papers should send their orders and remittances without delay to the Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. This will aid him to determine the number of copies to be printed for sale. Research Papers are not printed from electrotypes, and usually no more reprints can be had when the first and only printing is exhausted.

U. S. Department of Commerce National Bureau of Standards

RESEARCH PAPER RP1362

Part of Journal of Research of the National Bureau of Standards, Volume 26, February 1941

MICROSCOPIC STRUCTURE OF THE COTTON FIBER

By Charles W. Hock, Robert C. Ramsay, and Milton Harris [1]

ABSTRACT

During the examination of cotton fibers that had received various chemical treatments, a number of observations pertaining to the structural details of the fiber were made. Those phenomena which appeared to be new were investigated in detail. In addition, experiments described by earlier investigators were repeated in order to have a better composite picture of the structure of the fiber.

The cell wall of a cotton fiber consists of a primary and a secondary wall. The latter, which comprises the bulk of the fiber, consists of innumerable spirally oriented cellulose fibrils enclosed by a winding which also makes a spiral, but in the opposite direction from the former. Both the winding and the fibrils reverse their direction at frequent intervals along the axis of the fiber, their points of reversal being coincident. The secondary wall is enclosed by a thin primary wall. The latter is made up of fine crisscrossing strands of cellulose embedded in a membrane consisting principally of wax and pectic substance. The lumen also contains wax and pectic materials, plus various amounts of degenerated protoplasm.

When cotton fibers are swollen under certain conditions, a lamellate structure is discernible in the secondary wall. The number of these lamellae increases with the age of the fiber.

On treatment of cotton fibers with cuprammonium hydroxide reagent, the cellulose dissolves, leaving residues which vary in amount and in structure, depending upon the extent of purification of the fibers. The residue from raw and from dewaxed fibers consists of fragments of the primary wall, and of a lesser amount of material from the lumen. The behavior of the fibers in the reagent depends in part on their maturity. Immature fibers containing only small amounts of cellulose swell relatively little in the reagent, and the undissolved wax and pectic materials maintain the original tubular shape of the primary wall. When older fibers are given the same treatment, they swell abruptly, thereby causing the primary wall to break in many places, giving rise to "balloons."

Irregular swelling along the fiber axis, which results in the formation of balloons, appears to be dependent in part on the orientation of the fibrils, and in part on the constricting influences of the winding and of the primary wall.

CONTENTS

	Page
I. Introduction	94
II. Materials	94
III. Experimental procedure and results	95
1. Primary wall	95
2. Single fibers in cuprammonium reagent	96
3. Cross sections of fibers in cuprammonium reagent	97
4. Lamellae and growth rings	98
5. Fibrillar structure	100
6. Balloon formation	102
IV. References	104

[1] Research Associates at the National Bureau of Standards, representing the Textile Foundation.

I. INTRODUCTION

As a part of a program of research relating to the chemical and physical structure of textile fibers, microscopic examinations were made of cotton fibers which had received various chemical treatments. A number of observations pertaining to the structural details of the cell wall of the cotton fiber were made. Those phenomena which appeared to be new were investigated in considerable detail. In addition, experiments described by earlier investigators (Balls [1, 2, 3],[2] Denham [4], Sakostschikoff and Korsheniovsky [5], Kerr [6], Anderson and Kerr [7], Farr [8], and Barrows [9]) were repeated, and are reported here for the sake of completeness.

II. MATERIALS

Two varieties of *Gossypium hirsutum* L., Missdel-7 and Mexican Big Boll, were used in the investigation. Samples of each were furnished by the Agricultural Marketing Service of the U. S. Department of Agriculture. Missdel-7 was grown at the Delta Experiment Station, Stoneville, Miss. Only fibers from open bolls of this variety were used. The samples of Mexican Big Boll variety, grown at the Experiment Station at Raleigh, N. C., were from special case-historied material collected for joint fiber structural studies by the Bureau of Plant Industry and the Agricultural Marketing Service. The samples of the latter were principally young or immature fibers from bolls of known age. Fibers of different ages were obtained by tagging the flowers on the day of blooming and then collecting the bolls on successive dates so as to give material of the desired age. Two series of immature Mexican Big Boll fibers were used. One, collected during the summer of 1939, consisted of dried fibers in which deposition of secondary cellulose had begun. The other, collected during the summer of 1937, consisted of very young fibers with primary wall only. The second set of fibers was obtained from bolls, the locks of which, at the time of collection, had been placed in boiling water for a few minutes to untangle the fibers, and then stored in approximately 70-percent ethyl alcohol.

Those fibers which received no chemical treatments are termed "raw fibers." In the present paper the young Mexican Big Boll fibers which had been boiled in water and then placed in alcohol are also included in this category, although it is recognized that the treatment was not without effect on the composition of the fibers. The natural waxes were removed from some of the raw fibers by extraction with alcohol and ether for 24 hours each. These fibers are termed "dewaxed fibers." A portion of the dewaxed material was further purified by extraction with alkali according to the method recommended for the preparation of standard cellulose. The procedure was essentially the same as that described by Corey and Gray [10], except that the apparatus of Worner and Mease [11] was employed. These fibers, from which both wax and pectic substance were removed [12], are termed "depectinized fibers."

Cuprammonium hydroxide solution was prepared according to the recommendation of Mease [13]. The solution contained 240 ± 5 g of ammonia (NH_3), 15 ± 0.1 g of copper, and 1.0 g of sucrose per liter. The concentration of nitrite was less than 0.5 percent. When

[2] Figures in brackets indicate the literature references at the end of this paper.

not in actual use, the cuprammonium hydroxide was kept under nitrogen in tightly stoppered bottles in the refrigerator.

Trimethylbenzylammonium hydroxide was used in some of the experiments. The stock solution, which was 2.5 *N*, was diluted with water just before using.

For staining cellulose, iodine and sulfuric acid [14], zinc chloride–iodine [14], and an alkaline solution of Congo red (0.5 percent of dye in a 0.5-percent solution of sodium hydroxide) were used. Aqueous solutions of ruthenium red were used for staining pectic compounds [15]. Although ruthenium red is not a specific stain for pectic substance, native pectic compounds invariably take a red color in the presence of this dye [7].

III. EXPERIMENTAL PROCEDURE AND RESULTS

The cell wall of the cotton fiber is composed of a primary and a secondary wall. The primary wall, in accordance with the description of Anderson and Kerr [7], is considered to be the outer sheath of the fiber, or the portion of the wall formed as the fiber increases in length. Secondary wall then refers to that part of the wall which is laid down after the fibers cease to elongate appreciably. Only the thin primary wall encloses the protoplasm of the cotton hair as the fiber elongates during the first 13 to 20 days after its origin. Thereafter, the thickness of the wall is increased by a deposition of cellulose which comprises the secondary wall. For convenience in discussion, these two distinct layers are considered separately.

1. PRIMARY WALL

Young fibers, 3, 5, 8, and 15 days old, were mounted in water and examined microscopically. With ordinary light there was no evidence of structure in the thin primary wall. The walls stained deeply and uniformly with ruthenium red. When treated with an alkaline solution of Congo red, zinc chloride–iodine, or iodine and sulfuric acid, the fibers gave a faint color reaction. In older primary walls, for example in a 15-day fiber, an ill-defined pattern was observable in the wall.

On examination with crossed nicols, unstained fibers showed practically no birefringence. Insertion of a selenite plate between the nicols produced faint colors indicative of a predominantly transverse orientation. Although these fibers showed little birefringence when unstained, they were clearly anisotropic after treatment with a dye such as Congo red. This increase in birefringence upon staining is due, apparently, to the double refraction of the preferentially oriented dye molecules [16, 17].

When wax and pectic substance were removed from 15-day fibers the material which remained still had the outline of the original fiber and stained with Congo red, zinc chloride–iodine, and iodine and sulfuric acid. Between crossed nicols, the stained depectinized fibers were birefringent and showed a pattern like that in the stained walls of raw fibers. When the purified fibers were placed in cuprammonium hydroxide solution, they appeared to dissolve completely in the reagent.

The stained walls of both the raw and the depectinized fibers were suitable for observing the cellulose [3] of the primary wall. In agree-

[3] While the identity of this material has not been definitely established, X-ray investigations, staining reactions, and behavior in cuprammonium hydroxide solutions indicate that it is probably cellulosic in nature.

ment with the description by Balls [2, 3], and by Anderson and Kerr [7], the cellulose, as observed with crossed nicols, was found to be present as criss-crossing strands. At 45° with respect to the planes of the nicols, the fiber showed its maximum brightness and presented a reticulate appearance (fig. 1, *A*). When a fiber was rotated to the right or left of its position of maximum brightness, either one or the other of the systems of crisscrossing strands predominated the image (fig. 1, *B* and *C*). Measurements [4] of the angles which these birefringent strands made with the axis of the fiber were made on fibers of different ages. The results indicate that the angle of orientation decreases with increase in age of the fiber. In other words, the orientation was more nearly transverse in the younger than in the older ones. Likewise, the netlike structure of the cellulose, when the fibers were placed at 45°, appeared coarser (compare fig. 1, *A*, with fig. 2, *A*) and showed greater birefringence in the older fibers. The greater birefringence of the older fibers is not readily apparent from the photographs, as they were taken primarily to show the netlike structure. Measurements also indicated that the angle which the cellulose strands made with the fiber axis was greater at the tip than at the base. These differences were especially noticeable in 15-day fibers, where the growing cell had nearly attained its maximum length.

The above observations were confirmed by the examination of single walls obtained by cutting longitudinal sections of the fibers with the aid of a freezing microtome. Staining reactions and birefringence measurements gave essentially the same results as those obtained with intact fibers (fig. 2, *B*).

After the initiation of secondary thickening, the predominantly transverse orientation (with respect to the fiber axis) of the cellulose in the primary wall was masked, after a few days, by the more nearly parallel orientation of the cellulose in the secondary wall. This change was evident upon examination of the fibers with crossed nicols and a selenite plate, as well as by observation with ordinary light, and is in agreement with the results of the X-ray work of Berkley [18].

2. SINGLE FIBERS IN CUPRAMMONIUM REAGENT

The bulk of the cellulose of a cotton fiber is in the secondary wall. For this reason the attention of most investigators has been directed mainly to this part of the fiber.

In a previous investigation [19] it was found that the cellulose dissolved during the treatment of cotton with cuprammonium hydroxide solutions, leaving residues which varied in amount and in structure, depending upon the extent of purification of the fibers. When *raw* and *dewaxed* mature fibers were placed on microscope slides and then treated with cuprammonium reagent, the fibers immediately began to swell and twist, often forming balloon-like structures (fig. 3, *A*). The residue which remained after dissolution of the cellulose consisted principally of fragments (fig. 3, *B*) of the primary wall and, to a lesser extent, of material from the lumen. These residues were isotropic, contained pectic substance [12], and stained deeply with ruthenium red. They were often clumped into accordion-like structures which could be stretched with the needles of a micromanipulator to form continuous tubes corresponding to the condition in the intact

[4] The angle which the strands made with the axis of the fiber were measured at the center of the fiber by means of a rotating stage marked in degrees.

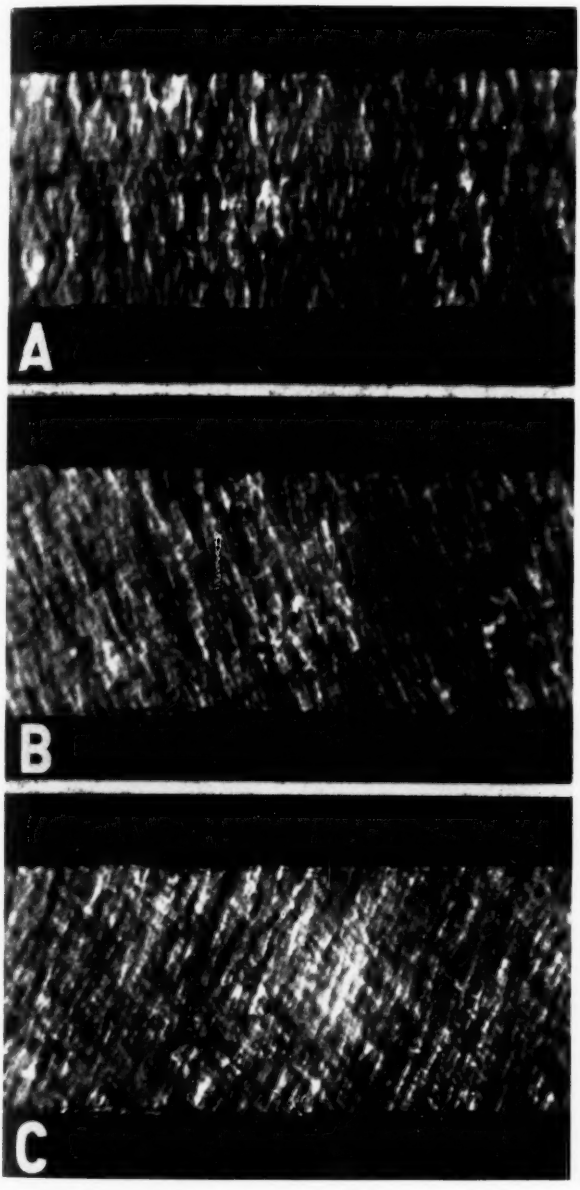


FIGURE 1.—*Raw 15-day cotton fiber stained with Congo red and photographed between crossed nicols.*

A, Fiber at 45° with respect to planes of the nicols, showing criss-crossing strands of cellulose in the primary wall; *B*, same fiber placed at about 15° to the plane of light passing through the polarizer, showing one system of strands which predominate in this position; *C*, same as *B*, with fiber placed at about 15° to the plane of light passing through the analyzer, and showing the other set of cellulose strands. Magnification ×800.

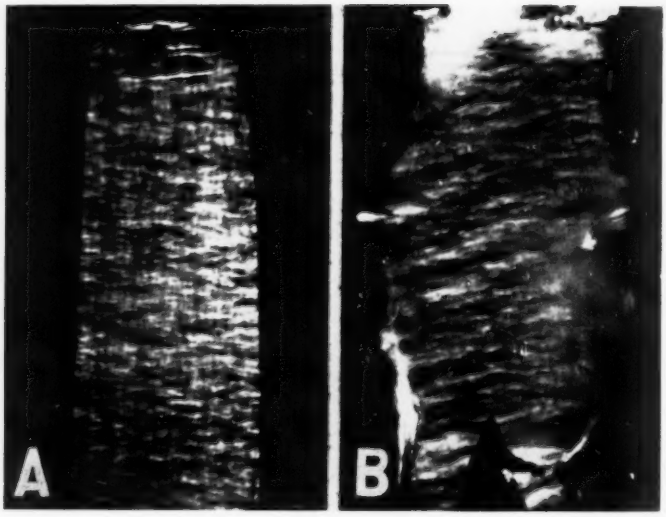


FIGURE 2.—*Raw fibers stained with Congo red and photographed between crossed nicols.*

A, 3-day fiber at 45° with respect to the planes of the nicols, showing strands of cellulose in the primary wall; *B*, single wall of a 15-day fiber, showing strands of cellulose in the primary wall. Fiber placed at 45° angle with respect to the planes of the nicols. Magnification ×800.

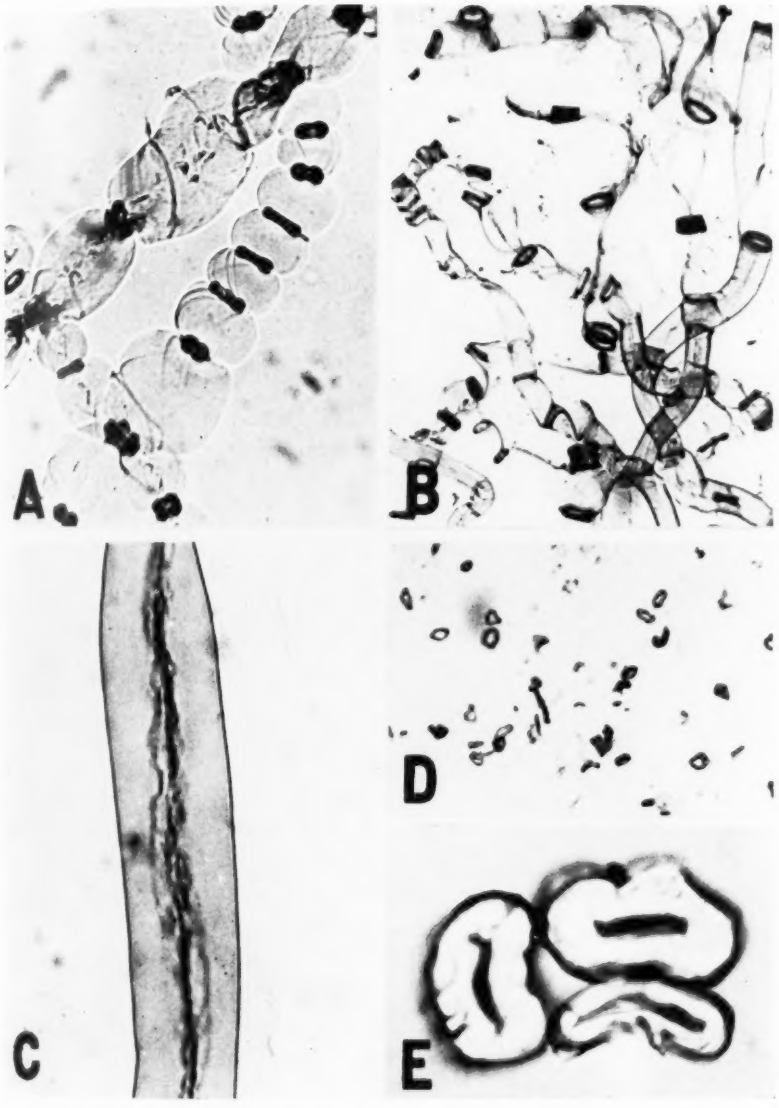


FIGURE 3.—*Raw cotton fibers in cuprammonium hydroxide solution.*

A, Mature fibers swollen to form balloons, magnification ×115; *B*, residue which remains after the cellulose of raw mature fibers has dissolved in the reagent, magnification ×115. *C*, typical residue of a 22-day raw fiber after treatment with cuprammonium reagent, magnification ×480; *D*, residue which remains after treating cross sections of mature raw fibers with cuprammonium reagent, magnification ×115; *E*, raw cross sections of mature cotton fibers swollen slightly in dilute cuprammonium hydroxide solution, showing deeply stained material in the lumen and at the periphery of the fibers, magnification ×1170.

All the fibers were stained with rutheniem red.

UNIV. OF MICH.

fibers. These tubes, which appeared to be soft and pliable, were readily torn by this treatment. *Depectinized* fibers dissolved in cuprammonium reagent, leaving only a negligible amount of isotropic residue which exhibited no definite cell structure.

When immature fibers in which only a small amount of secondary thickening had occurred were treated with cuprammonium reagent, relatively few visible changes took place. Few "balloons" were formed and, therefore, the primary wall was left practically intact. The mild reaction of young fibers in cuprammonium hydroxide was in sharp contrast to the swelling, writhing, and twisting of mature fibers. Young fibers containing only small amounts of cellulose swelled little in the reagent, and the undissolved wax and pectic materials maintained the original tubular shape (fig. 3, *C*) of the primary wall. By noting their behavior in cuprammonium hydroxide, it was thus possible to distinguish mature from immature fibers.

3. CROSS SECTIONS OF FIBERS IN CUPRAMMONIUM REAGENT

In order to study further the behavior of cotton fibers in cuprammonium reagent, cross sections of *raw*, *dewaxed*, and *depectinized* fibers were made. The usual procedure for making sections, whereby the fibers are embedded in paraffin, collodion, gelatin, or a similar material, could not be used, as it is necessary to remove the embedding materials from the sections after they are cut, and the solvents used for this purpose also take out some of the waxes or pectic substance from the untreated cotton fiber. Accordingly, cross sections were made by simpler methods, which may not give sections as uniform as those obtained with embedded material, but have the advantage of leaving the fibers unaltered in their chemical composition.

By means of the device of Hardy [20], fibers were cut into sections about 10 microns thick without embedding. The sections were placed on microscope slides and then examined, as cuprammonium hydroxide solution was drawn under the cover glass. The sections of *raw fibers* swelled in the reagent, and in a short time the cellulose appeared to be dissolved, leaving a residue in the form of rings and other fragments of material (fig. 3, *D*) which stained with ruthenium red. Most of the residue came from the surface of the fiber, a lesser amount from the lumen. When this behavior was observed with crossed nicols, it was noted that the birefringence of the sections was lost as the cellulose dissolved and that the residue was isotropic. When sections were swollen only slightly in dilute cuprammonium reagent and then placed in ruthenium red, deeply stained red bands were discernible at the circumference and in the lumen (fig. 3, *E*). Sections of *dewaxed fibers* behaved in essentially the same way as raw cotton. Dissolution of the cellulose took place rapidly, leaving an isotropic residue which stained with ruthenium red and was similar to that obtained from the raw fibers. When sections of *depectinized fibers* were treated with cuprammonium hydroxide they dissolved almost completely in the reagent.

Cuprammonium reagent dissolves the cellulose of single cotton fibers or of cross sections in a relatively short time. On the other hand, dilute solutions of the reagent swell the fibers but dissolution of the cellulose is greatly retarded. Swelling fibers in this way offers an excellent means of observing details of their structure. Cross sections about 10 microns thick were placed on microscope slides and then

treated with cuprammonium hydroxide solution which had been diluted four to six times with concentrated ammonium hydroxide. The sections swelled to many times their original diameter, thereby revealing a lamellate structure (fig. 4, *A* and *B*). During the course of swelling the sections usually inverted, so that the region originally at the surface of the fiber occupied the former position of the lumen. Where inversion was incomplete, cone-shaped structures were formed (fig. 4, *C*). Inversion was observed in sections of fibers from which the wax and pectic substance had been removed, as well as in raw fibers. Very thin sections of mature fibers sometimes swelled without inverting. It should be pointed out, however, that a slight tear at the periphery, such as might easily occur in making sections only a few microns in thickness, would allow the sections to expand without inverting. Cross sections of immature fibers with only a small deposit of cellulose usually did not turn inside out. Apparently in immature fibers the lumen is sufficiently large to allow for expansion of the relatively small amount of cellulose. The regularity with which inversion accompanies the swelling of sections of mature fibers indicates that structural features of the fiber itself may be responsible for this phenomenon. Since inversion occurs regularly, even in depectinized fibers, the insoluble wax and pectic substance in the primary wall cannot be primarily responsible for the observed behavior. It appears, therefore, that inversion may be due to a constricting influence, possibly of a layer or membrane which has not yet been identified, or what seems to be more probable, to a differential rate of swelling between the inner and outer layers of the cellulose of the secondary wall.

4. LAMELLAE AND GROWTH RINGS

Lamellae were not readily seen in untreated cross sections but were clearly visible in properly swollen material. When treated with cuprammonium reagent diluted four to six times with concentrated ammonium hydroxide, swelling of the cellulose took place in such a way as to reveal lamellae. Slight pressure on the cover glass helped to spread the swollen sections and thus render the lamellae more distinct, but this technique was not necessary for making these structures visible. When, for example, the cover glass was supported by bits of broken cover glass to prevent pressure on the sections, swelling took place as usual and lamellae were clearly discernible. Likewise, when the sections were swollen in a hanging drop,[5] lamellae were similarly observed.

On treating the swollen sections with Congo red, alternating lamellae appeared to be deeply and lightly stained with the dye (fig. 4, *A* and *B*). Between crossed nicols these sections showed alternating layers of strong and of weakly birefringent material (fig. 4, *D*).

The above-described techniques were applied to Missdel-7 and to Mexican Big Boll cottons. In both varieties the structure of the fibers was essentially the same. The average number of lamellae in a mature fiber was between 45 and 50. Individual lamellae varied in thickness from 0.1 to 0.2 micron.[6] The first lamella appeared to be deposited between the 18th and 20th day after flowering, the

[5] The sections were placed in a drop of dilute cuprammonium hydroxide solution on a cover glass, and the latter was inverted over a depression in a microscope slide.

[6] The width of the lamellae was obtained by dividing the thickness of the wall of the unswollen fiber by the number of layers counted after swelling.

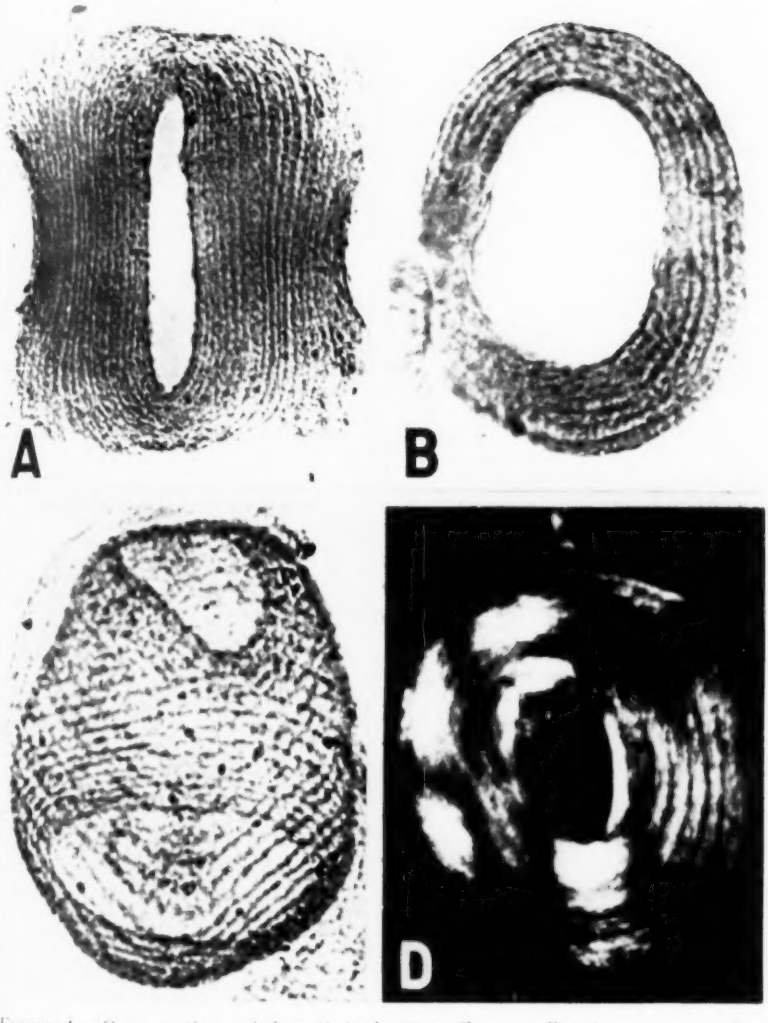


FIGURE 4.—*Cross sections of depectinized cotton fibers swollen in cuprammonium hydroxide solution to show lamellae.*

A, Section of mature fiber, showing lamellae, magnification ×400; *B*, section of immature (25-day) fiber to show lamellae, magnification ×910; *C*, section of mature fiber swollen to show lamellae, and partly inverted to produce a cone-shaped structure; magnification ×400; *D*, section of 26-day fiber to show lamellae. Photographed between crossed nicols, magnification ×670.

All the fibers were stained with Congo red.

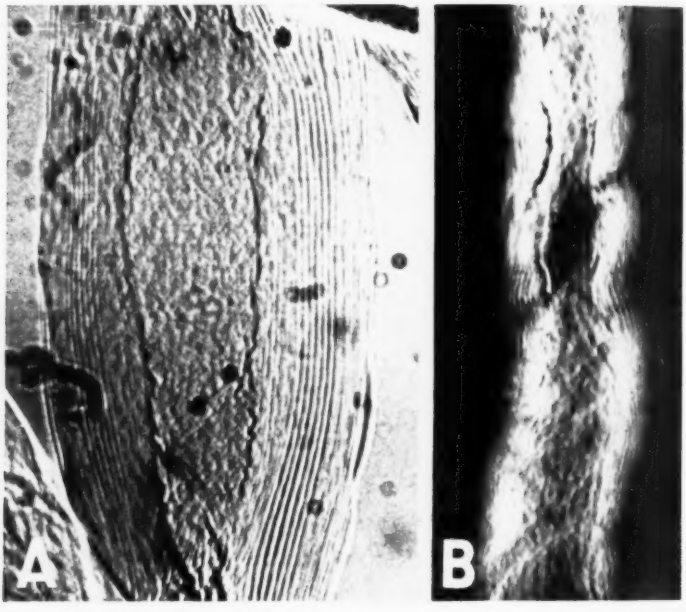


FIGURE 5.—*Depectinized cotton fibers swollen in cuprammonium hydroxide solution to show striations.*

A, 51-day fiber, showing striations, magnification ×500; *B*, mature fiber, showing striations. Photographed between crossed nicols, magnification ×420.
All the fibers were stained with Congo red.

number increasing thereafter as the fiber approached maturity. The increase in the number of lamellae with age can be seen by comparing figure 4 (*A*) with figure 4 (*B*).

The lamellae, so clearly observed in swollen cross sections, can also be seen in longitudinal view [1], when the microscope is focused midway between the upper and lower surfaces of a swollen fiber. Fibers from which both wax and pectic substance were removed were preferable for these observations, as the purified fibers swelled more uniformly than raw fibers. When observed in this way, the fibers showed striations [7] running parallel to the fiber axis and extending from the lumen to the outer surface (fig. 5, *A*). The striations were readily discernible in unstained mounts, but the addition of Congo red rendered the pattern more distinct. When examined with crossed nicols alternating stripes of strongly and of weakly birefringent material were observed (fig. 5, *B*). As the fiber matures, the thickness of the wall becomes greater and the number of striations is increased. The number of striations in Mexican Big Boll from 18 days to maturity was found to agree with the number of lamellae counted in cross sections.

According to Kerr [6], two lamellae, one compact and one porous, are deposited each day during the period of secondary-wall deposition. These two lamellae (or two adjacent striations) together constitute a daily growth ring. Using "growth ring" in this sense, the relationship between these rings and the age of the fiber is shown graphically in figure 6. It is readily seen that there is a close correlation between the counts of lamellae and of striations. For the first 10 to 15 days after the initiation of secondary deposition, one growth ring was laid down per day. Thereafter, until maturity, additional growth rings were formed, but the rate of one ring per day was apparently not maintained. However, closer examination of the data showed that the counts of lamellae and of striations varied considerably in fibers older than 35 days. This may mean that individual fibers continued to deposit one growth ring daily, but that the average for the group is lowered because of the early maturation of some of the fibers.

These results confirm the work of Balls [1] and Kerr [6]. They are not in agreement with the results of the investigations by Sakostchikoff and Korsheniovsky [5] and Barrows [9], who were unable to establish a consistent relationship between the number of lamellae and the age of the fibers.

Figure 7 shows the increase in width of the secondary wall with the age of the cotton fiber. After about the first 15 days of secondary-wall formation there is only a slight increase in thickness as the fibers continue to mature. This is consistent with a recent investigation [21] in which it was demonstrated that the percentage of cellulose in a series of cotton samples of different ages increased rapidly until about the 35th day after flowering, and that thereafter there was only a small further increase with age. A comparison of the curves of figures 6 and 7 suggests that the initial growth rings are wider than those laid down after the 35th day.

[7] In the present paper, the layers observed in cross section are called lamellae, whereas the layers observed in longitudinal view are called striations. Both lamellae and striations refer, however, to the same structures.

Lamellae, or striations, were also revealed when fibers were treated with reagents other than cuprammonium hydroxide. Trimethylbenzylammonium hydroxide had essentially the same effect as cuprammonium reagent. Combined treatment with sodium hydroxide and carbon disulfide, and treatment with phosphoric acid also

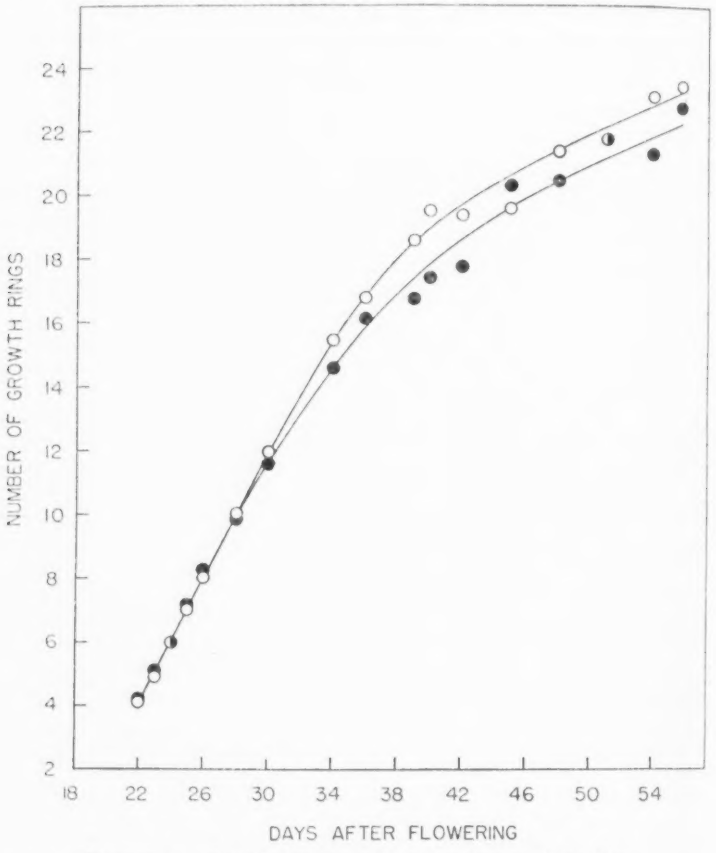


FIGURE 6.—*Growth rings in relation to the age of the fiber.*

Open circles represent growth rings estimated from the number of lamellae counted in fiber cross sections (25 sections counted for each point of the curve). Solid circles represent growth rings estimated from the number of striations counted in longitudinal view (50 sections counted for each point of the curve).

revealed a lamellate structure in the fiber, but cuprammonium hydroxide and trimethylbenzylammonium hydroxide were preferable for this work.

5. FIBRILLAR STRUCTURE

The investigation of the structural details of the cotton fiber was continued with single fibers. These were mounted in distilled water (fig. 8, *A*) and then examined with crossed nicols. When the fibers were placed approximately parallel to the plane of vibration of light through one of the prisms, the high birefringence of the fiber was interrupted by dark extinction bands at irregular intervals along its

axis (fig. 8, *B*). The same phenomenon has previously been noted by Balls [2]. When a selenite plate was inserted between the nicols, contrasting colors, indicative of differences in orientation, appeared in regions adjacent to the extinction bands (fig. 8, *C*). The number of convolutions or twists of the fiber which occurred in a region

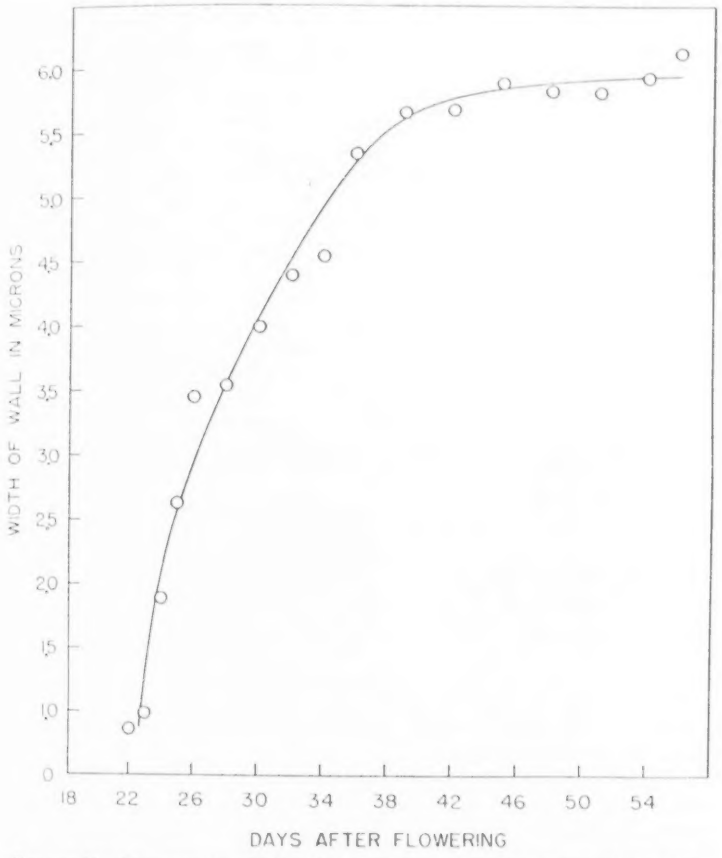


FIGURE 7.—*Increase in width of the wall with increase in age of the fiber.*
Each point of the curve represents an average of 200 measurements.

between two bands of extinction varied considerably, depending apparently on the length of the region. However, in a region which showed uniform optical properties the convolutions were always in the same direction.

The optical variations just described were also correlated with differences in structure observed in depectinized [8] fibers swollen in dilute cuprammonium hydroxide or in solutions of trimethylbenzylammonium hydroxide. For this experiment the cuprammonium reagent was diluted three to four times with concentrated ammonia, and the trimethylbenzylammonium hydroxide was diluted with

[8] Depectinized fibers were best suited for these experiments, since, after removal of the outer membrane of wax and pectic substance, irregular swelling along the axis could be minimized.

water to a concentration of 2.2 to 2.4 *N*. Upon swelling, the fibers clearly revealed a fibrillar structure. The bulk of a fiber was made up of innumerable fine fibrils oriented at an acute angle with respect to the axis of the fiber. The fibrils were enclosed by a winding [9] which appeared to be a little coarser than the fibrils themselves.

In fibers which had been swollen only slightly, it was observed that this outer winding was closely spaced and made a moderately steep spiral (fig. 9, *A* and *B*). The winding, like the bulk of the fiber, stained with Congo red, zinc chloride–iodine, and iodine and sulfuric acid, and appeared to dissolve in cuprammonium reagent. It made either an S or a Z twist [10] around the fiber, reversal of direction taking place (fig. 9, *C*) many times in a single fiber. Examination showed that both types of twist occur with about equal frequency. Invariably the orientation of the fibrils immediately beneath the outer winding was in the opposite direction. In other words, if the winding made an S twist, the fibrils immediately beneath made a Z twist. As far as could be determined, all of the fibrils under the winding at any one place appeared to run in the same direction.

It became apparent that reversals of direction of the cellulose fibrils and of the winding were responsible for the optical differences observed with crossed nicols, and that the band of extinction was the place at which the reversals occurred. The number of optical reversals occurring in a single fiber or fragment invariably was the same as the number of fibril reversals. Likewise, the relationship between the fibrillar orientation and the convolutions became clear. It was found that the direction of a convolution usually was the same as the spiral direction of the winding. That is to say, when the winding made an S twist around the axis, the convolutions in that region also showed an S twist.

The fibrillar structure of the secondary wall was studied to better advantage upon handling the fibers with microneedles. In some cases the outer winding was picked up by the needles and pulled away from the rest of the fiber or the winding was pulled along the axis of the fiber in such a way as to form a clump. Needles were brought into the innumerable fine fibrils comprising the bulk of the cellulose of the secondary wall only by careful manipulation. The difficulty in picking up these fibrils appeared to be due to lateral cohesion of the fibrils themselves. After insertion of the needles, however, the fibrils could be separated from one another (fig. 10, *A* and *B*) and otherwise subjected to micromanipulative methods.

6. BALLOON FORMATION

Depectinized fibers were also examined in cuprammonium hydroxide solutions of various concentrations. It soon became clear that there was a relationship between the optical differences and fibril reversals on the one hand, and balloon formation on the other. In *concentrated* cuprammonium hydroxide solutions (reagent diluted not more than three times with concentrated ammonia) dissolution of the cellulose of the fibers took place in a relatively short time. The fibers appeared to break into large chunks which became progressively smaller as dis-

[9] Although the winding closely resembles the fibrils which make up the bulk of the fiber, to facilitate discussion it will be considered separately.

[10] According to ASTM Standards on Textile Materials [22], a yarn or cord has an S twist if, when held in a vertical position, the spiral conforms in slope to the central portion of the letter S, and a Z twist if the spirals conform to the central portion of the letter Z. In the present paper, the same terminology will be applied to fibrils.

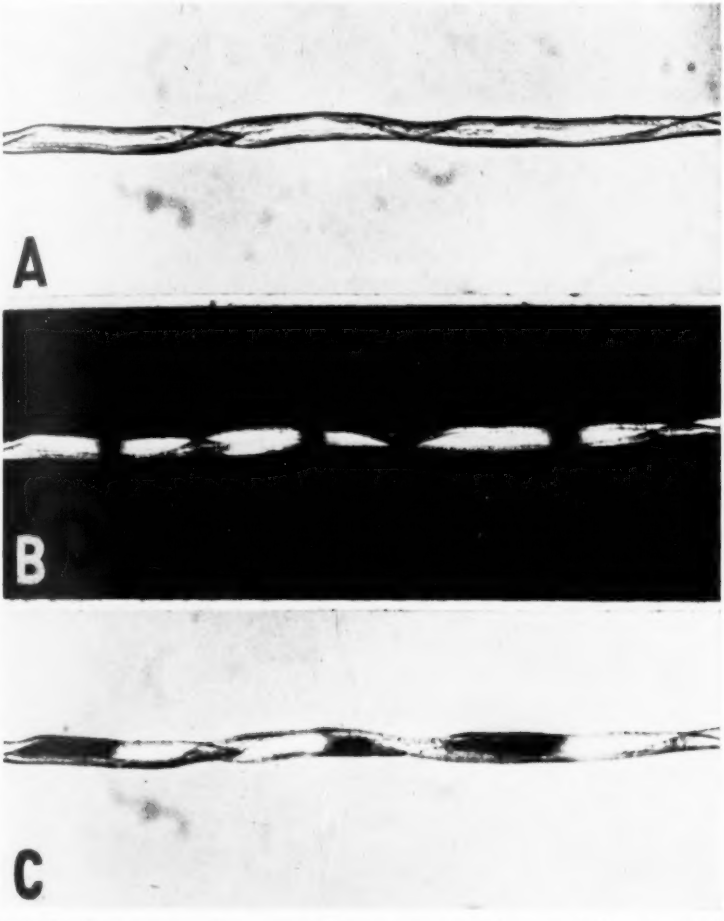


FIGURE 8.—*Raw mature fiber mounted in water and photographed under various lighting conditions.*

A, Ordinary light, showing convolutions in the fiber; *B*, between crossed nicols, showing bands of extinction where the fibrillar orientation reverses; *C*, between crossed nicols and with selenite plate, showing the effect due to differences in color in regions adjacent to the extinction bands: magnification, *A*, *B*, and *C*, ×180.

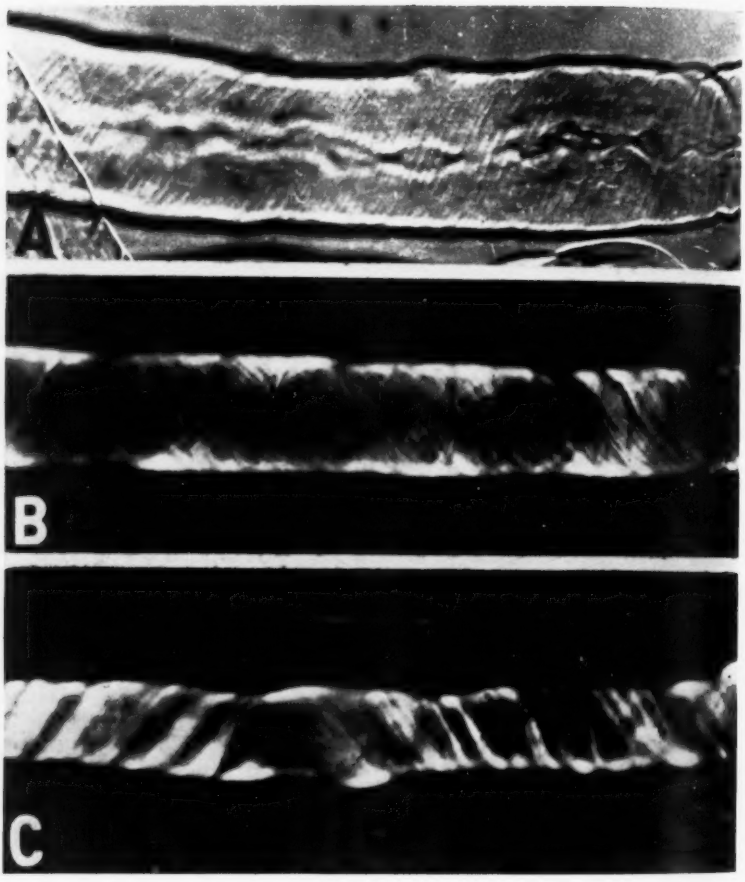


FIGURE 9.—*Depectinized fibers showing the winding.*

A, Fiber swollen slightly in dilute cuprammonium hydroxide solution; *B*, fiber treated with zinc chloride-iodine and photographed between crossed nicols; *C*, fiber, showing reversal, treated with zinc chloride-iodine and photographed between crossed nicols; magnification, *A*, *B*, and *C*, ×420.

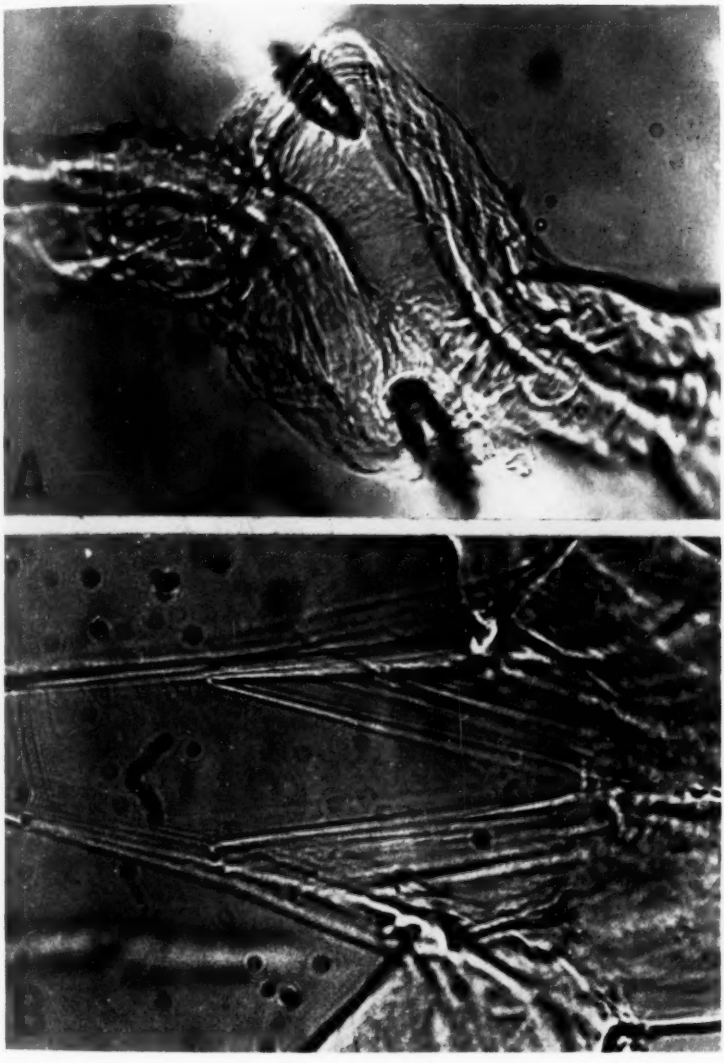


FIGURE 10.—*Depectinized fibers swollen in dilute cuprammonium hydroxide solution and handled with microneedles.*

A, A single fiber being dissected with microneedles to show fibrillar structure; *B*, fibrils being pulled away from the fiber. (The needles are not shown in the photomicrograph.) Magnification, *A* and *B*, ×500.

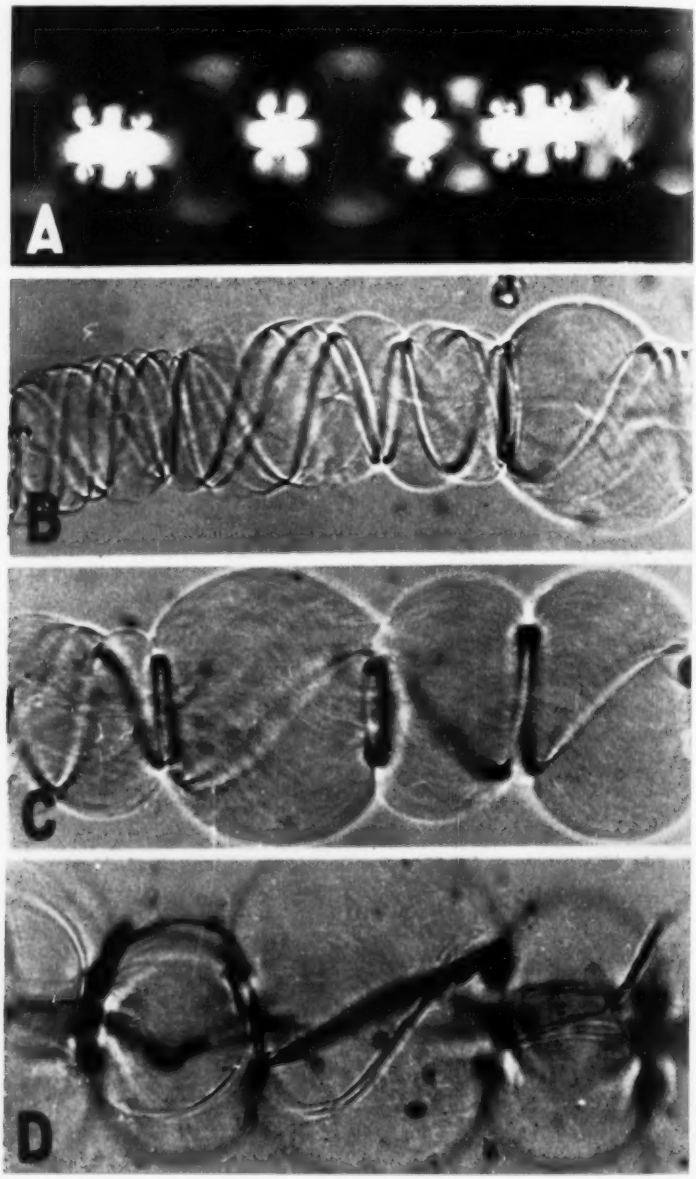


FIGURE 11.—*Cotton fibers swollen in cuprammonium hydroxide solution to produce balloons.*

A, Depectinized fibers, showing balloons, photographed between crossed nicols, (selenite plate inserted), magnification ×260; *B*, depectinized fibers, showing winding. The coarse appearance of the winding may have resulted from the coalescence of individual fibrils during incipient balloon formation, magnification ×420. *C*, typical string of balloons in a depectinized fiber, magnification ×500; *D*, string of balloons in a raw fiber, showing the winding and fragments of the primary wall connecting the constrictions. Fiber stained with ruthenium red. Magnification ×440.

solution proceeded. Observed between crossed nicols, the highly birefringent fibers were seen to break up into fragments which steadily became less birefringent until the field of the microscope was black. In *dilute* solutions (diluted between 3 and 10 times) the fibers swelled greatly but complete dissolution generally did not occur. As swelling took place balloons were occasionally formed. When the behavior of the fibers was observed with crossed nicols, the fibers were seen to become progressively dimmer as the cellulose dissolved. The expanded part of a balloon was the first to lose its birefringence, since here the cellulose had become highly disorganized. On the other hand, the constrictions were slow in losing their anisotropy (fig. 11, *A*), since the reagent penetrated these regions slowly. When fibers which had lost much of their birefringence during swelling in cuprammonium solution were placed in water, some of the original birefringence of the fibers, as well as some of their fibrillar appearance, was restored.

The occurrence of balloons in *depectinized* fibers was somewhat surprising, since experiments by ourselves and by others indicated that the primary wall plays an important role in the formation of balloons. Accordingly, the phenomenon was studied further, especially in depectinized fibers from which the wax and pectic substance had been removed. The optimum concentration of cuprammonium hydroxide solution for observing balloons in these fibers was between 15 and 20 percent (1.5 to 2.0 ml of the reagent diluted to 10 ml by the addition of concentrated ammonia). When depectinized fibers were placed in these solutions they often swelled unevenly along the axis, thereby giving rise to balloons (fig. 11, *B* and *C*). The expanded portion of the balloons almost invariably occurred between the points at which the fibrils reversed, whereas the reversal points themselves formed the constrictions between adjacent balloons. Fragments of depectinized fibers which showed no extinction bands usually swelled without the formation of balloons, and the direction of the fibrils was found to be constant. Because of the orientation of the cellulose within the fiber, the greatest swelling occurs perpendicular to the fiber axis, while at the same time there is an actual decrease in fiber length. As these changes in length and in width take place, the winding assumes a more transverse position (fig. 11, *B*) and becomes clumped at the constrictions. However, the uniform lateral expansion of the fiber is hindered somewhat by this same winding which does not expand freely in that direction. Accordingly, the winding also appears to exert a constricting influence on the swelling of the fibers.

It has been recognized that the primary wall is responsible, at least in part, for the formation of balloons. The rupture of this wall into a series of constrictions which restrains the expanding cellulose has frequently been described. Furthermore, fragments of *raw* fibers which show uniform optical characteristics often produce many balloons as swelling occurs, whereas similar fragments of depectinized fibers do not produce balloons. When balloons are formed in raw fibers the cellulose fibrils may push through the winding and the primary wall, so that a constriction consists of a clump of the winding as well as of a portion of the primary wall (fig. 11, *D*). Thus the above experiments suggest that the irregular swelling along the fiber axis, which frequently results in balloon formation, is dependent in part on the orientation of the fibrils and in part on the constricting influences of the winding and of the primary wall.

IV. REFERENCES

[1] W. L. Balls, Proc. Roy. Soc. (London) [B] **90,** 342 (1919).
[2] W. L. Balls, Proc. Roy. Soc. (London) [B] **72,** 72 (1923).
[3] W. L. Balls, Studies of Quality in Cotton (Macmillan & Co. Ltd., London, 1928).
[4] H. J. Denham, J. Textile Inst. **14,** T86 (1923).
[5] A. P. Sakostschikoff and G. A. Korsheniovsky, Faserforschung **9,** 249 (1932).
[6] T. Kerr, Protoplasma, **27,** 229 (1937).
[7] D. B. Anderson and T. Kerr, Ind. Eng. Chem. **30,** 48 (1938).
[8] W. K. Farr, Contrib. Boyce Thompson Inst. **10,** 71 (1938).
[9] F. L. Barrows, Contrib. Boyce Thompson Inst. **11,** 161 (1940).
[10] A. B. Corey and H. L. Gray, Ind. Eng. Chem. **16,** 853, 1130, (1924).
[11] R. K. Worner and R. T. Mease, J. Research NBS **21,** 609 (1938) RP1146.
[12] R. L. Whistler, A. R. Martin, and M. Harris, J. Research NBS **24,** 555 (1940) RP1299.
[13] R. T. Mease, J. Research NBS **22,** 271 (1939) RP1179.
[14] C. Dorée, The Methods of Cellulose Chemistry, D. Van Nostrand Co., Inc., New York, N. Y., (1933).
[15] L. Mangin, Compt. rend. **116,** 653 (1893).
[16] D. R. Morey, Textile Research **3,** 325 (1933).
[17] A. Frey-Wyssling, Die Stoffausscheidung der höheren Pflanzen. Julius Springer, Berlin (1935).
[18] E. E. Berkley, Textile Research **9,** 355 (1939).
[19] C. W. Hock and M. Harris, J. Research NBS **24,** 743 (1940) RP1309.
[20] J. I. Hardy, U. S. Dept. of Agriculture Circular 378 (1935).
[21] R. L. Whistler, A. R. Martin, and I. M. Conrad, J. Research NBS **25,** 305 (1940) RP1326.
[22] ASTM Standards on Textile Materials, Report of Committee D–13, p. 5 (1940).

WASHINGTON, September 28, 1940.

U. S. DEPARTMENT OF COMMERCE NATIONAL BUREAU OF STANDARDS

RESEARCH PAPER RP1363

Part of Journal of Research of the National Bureau of Standards, Volume 26, February 1941

OBLIQUE-INCIDENCE RADIO TRANSMISSION AND THE LORENTZ POLARIZATION TERM

By Newbern Smith

ABSTRACT

The force on an electron in an ionized medium may be written $\mathscr{E}+4\pi aP$ per unit charge, where $\mathscr{E}$ is the electric force, P the polarization of the medium, and a a constant which is zero on the Sellmeyer theory and may have a value of 1/3 on the Lorentz theory. A well-defined distinction exists between maximum usable frequencies for ionospheric radio transmission calculated on the basis of the two theories. This paper describes a type of experiment for measuring maximum usable frequencies and comparing them with calculated values. The results indicate that a is probably zero, at least in these experiments. There is also described a step-by-step method of solving the virtual-height integral equation, necessary in much theoretical work concerning the ionosphere.

CONTENTS

	Page
I. Introduction	105
II. Theory, neglecting the earth's magnetic field	106
III. Effect of the earth's magnetic field	110
IV. Experimental observations	111
V. Conclusions	115
VI. References	116

I. INTRODUCTION

The question whether the force on an electron in an ionized medium should be taken as merely the applied electric force, $\mathscr{E}$, or the so-called Lorentz force, $\mathscr{E}+4\pi aP$ [1],[1] where P is the polarization of the medium and a is a constant which is usually taken as 1/3, is of fundamental importance in the theory of the ionosphere. The two theories are known, respectively, as the Sellmeyer theory and the Lorentz theory, and the difference between them represents a difference of 50 percent in the equivalent electron density in the ionosphere.

The argument had apparently been settled in favor of the Sellmeyer theory in a comprehensive theoretical paper by Darwin [2], when the whole question was reopened by a radio experiment of Booker and Berkner [3], which they were able to interpret only in terms of the Lorentz theory. In Australia, Martyn and Munro [4] explained the same kind of radio experiment in terms of the Sellmeyer theory, but Appleton, Farmer, and Ratcliffe [5] in England attacked this interpretation.

The radio experiments referred to were all performed at vertical incidence, in the neighborhood of the gyrofrequency, or precessional frequency of electrons in the earth's magnetic field (about 1.4 Mc/s). The purpose of this paper is to present the results of a different kind

[1] Figures in brackets indicate the literature references at the end of this paper.

of radio experiment—one involving ionospheric transmission of radio waves over a distance at oblique incidence, at a high frequency, and during the daylight hours.

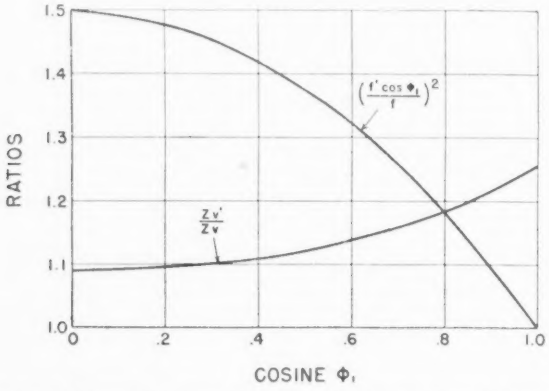


FIGURE 1.—*Oblique-incidence relations for Sellmeyer and Lorentz theories.*

z_v'=height of equivalent triangular path, $a=1/3$
z_v=height of equivalent triangular path, $a=0$ (equal to virtual height at equivalent vertical-incidence frequency)
f=equivalent vertical-incidence frequency, $a=1/3$
$f' \cos \phi_1$=equivalent vertical-incidence frequency, $a=0$
ϕ_1=angle of incidence of waves upon the ionized layers.

II. THEORY, NEGLECTING THE EARTH'S MAGNETIC FIELD

There exists a well-defined distinction in the behavior of oblique-incidence radio transmission according to the two theories. This distinction is that the maximum usable frequency for radio sky-wave transmission over a given distance is greater for the Lorentz theory than for the Sellmeyer theory.

In two previous papers [6, 7] the author outlined a method of calculating maximum usable frequencies and heights of reflection for oblique-incidence radio transmission directly from the curve of frequency versus virtual height obtained at vertical incidence. This curve will hereafter be called the (z_v,f) curve, where z_v is the virtual height and f is the vertical-incidence frequency. The virtual height is obtained by measuring the time, Δt, taken by a pulse to travel up to the ionosphere and back. It is given by

$$z_v=\frac{c\Delta t}{2}=\int_0^{z_0}\frac{d}{df}(\mu f)\,dz,$$

where c=velocity of light,
f=frequency of waves,
μ=refractive index of the ionosphere,
z_0=actual level of reflection, and
z=measured vertically upward.

The method of calculating maximum usable frequencies was based on the simple ray theory of a stratified isotropic ionosphere, whose refractive index, μ', for waves of frequency f' was given by

$$\mu'=\sqrt{1-\frac{f_0^2}{f'^2}} \tag{1}$$

where f' is the actual wave frequency and $f_0^2=Ne^2/\pi m$, N being the ionization density expressed as a function of height.

The calculation of maximum usable frequencies consists essentially in solving graphically the two equations $z_v=z_v(f)$, the (z_v,f) curve for vertical incidence, and $z_v=z_v(f, f', D)$, the transmission equation for oblique incidence, obtained below. Here D is the distance of transmission, f' the wave frequency, and f the equivalent vertical-incidence frequency, or frequency of the wave reflected, at vertical incidence, at the same level as is the wave frequency, f', over the given distance, D.

The transmission equation is found by combining the expression for the equivalent triangular path, obtained geometrically, with the expression for the equivalent vertical-incidence frequency, f, and the relation between z_v and z_v', the height of the equivalent triangular path.

The intersection of the graphs of the two equations gives both the equivalent vertical-incidence frequency, f, and the virtual height of reflection corresponding to transmission of the given wave frequency, f', over the given distance, D. The greatest value of f' for which a real solution exists is the maximum usable frequency over the given distance.

When the Lorentz polarization term a is included, the refractive index, μ', is given [8] by

$$\mu'=\sqrt{1-\frac{f_0^2}{f'^2+af_0^2}}. \tag{2}$$

The virtual height at the vertical-incidence frequency, f, is

$$z_v=\int_0^{z_0}\frac{1-a(1-\mu^2)^2}{\mu}\,dz, \tag{3}$$

where z_0 is the level of reflection, and μ is μ' with f substituted for f'. The equivalent vertical-incidence frequency, f, is given by

$$f=f'\cos\phi_1\sqrt{\frac{1-a}{1-a\cos^2\phi_1}}, \tag{4}$$

where ϕ_1 is the angle of incidence of the actual wave upon the ionosphere.

The height, z_v', of the equivalent triangular path may be calculated as follows. Let ϕ be the angle between the ray path and the vertical at the height z. Then

$$z_v' = \cos\,\phi_1 \int_0^{z_0} \frac{dz}{\mu' \cos\,\phi} = \cos\,\phi_1 \int_0^{z_0} \frac{dz}{\sqrt{\cos^2\,\phi_1 - \frac{f_0^2}{f'^2 + a f_0^2}}}.$$

This becomes, on substituting f for f',

$$z_v' = \int_0^{z_0} \sqrt{\frac{f^2 + a\frac{(1-a)\,\cos^2\,\phi_1 f_0^2}{1-a\,\cos^2\,\phi_1}}{f^2-(1-a)f_0^2}}\,dz. \tag{5}$$

To obtain the transmission equation, we get geometrically the value of

$$z_v' = z_v'(D, \phi_1).$$

We then insert in eq 6 the value of ϕ_1 obtained from eq 4 in terms of f and f', and multiply z_v' by the ratio $z_v : z_v'$ determined by comparing eq 3 and 5.

This analysis has been made for a flat ionosphere, but may be extended to a curved ionosphere in the manner indicated in reference [7].

For $a=1/3$, the ratio f'/f is greater than for $a=0$, and so the wave frequency, f', corresponding to a given f and z_v is in general greater. If z_v were equal to z_v', therefore, the calculated maximum usable frequency would be much greater for $a=1/3$ than for $a=0$, the difference increasing with the angle of incidence, ϕ_1. We must, however, examine the relation between z_v and z_v'. Now z is a function of f_0^2, so if we put $x \equiv f_0^2/f^2$ we may rewrite eq 3 and 5:

$$z_v = \int_0^{\frac{1}{1-a}} \frac{1-a\frac{x^2}{(1+ax)^2}}{\sqrt{1-\frac{x}{1+ax}}} \frac{dz}{dx}dx. \tag{7}$$

$$z_v' = \int_0^{\frac{1}{1-a}} \sqrt{\frac{1+a\frac{(1-a)\,\cos^2\,\phi_1}{1-a\,\cos^2\,\phi_1}x}{1-(1-a)x}}\,\frac{dz}{dx}dx. \tag{8}$$

Equation 7 gives z_v as a function of f^2 and eq 8 gives z_v' as a function of f^2 and $\cos\,\phi_1$, so that z_v/z_v' may be expressed as a function of f^2 and $\cos\,\phi_1$.

The ratio z_v/z_v' is, for $a=1/3$, in general less than unity, that is, less than the value it had for $a=0$. This tends to decrease the calculated maximum usable frequencies. To what extent the decrease due to this effect compensates for the increase due to the greater ratio of f'/f must be determined in individual cases.

Let us first consider the case of a linear gradient of ionization density, where $z=kf_0^2$. Here $dz/dx=kf^2$ and the determination of z_v and z_v' is easy. We obtain

$$z_v = 3z_0\left(\frac{1-a}{a}\right)\left[-1+\frac{1}{\sqrt{a(1-a)}}\,\tan^{-1}\sqrt{\frac{a}{1-a}}\right] \tag{9}$$

$$z_v' = z_0\left[1+\frac{1}{\sqrt{a\,\cos^2\,\phi_1(1-a\,\cos^2\,\phi_1)}}\,\tan^{-1}\sqrt{\frac{a\,\cos^2\,\phi_1}{1-a\,\cos^2\,\phi_1}}\right]. \tag{10}$$

Figure 1 shows the ratio of z_v'/z_v for $a=1/3$, which is the same as the ratio of z_v' for $a=1/3$ to z_v' for $a=0$, plotted against cos ϕ_1. It shows also the ratio $(f'\cos\phi_1/f)^2$ for $a=1/3$, which is the same as the square of the ratio of f for $a=0$ to f for $a=1/3$, plotted against cos ϕ_1. The two ratios are equal at about cos $\phi_1=0.8$. This means that, for the distance corresponding to this value of cos ϕ_1, the wave frequency corresponding to reflection at a point on the linear part of a (z_v,f^2) curve is the same for $a=0$ as for $a=1/3$. At shorter distances the wave frequency is less, and at greater distances it is greater, for $a=1/3$ than for $a=0$.

If the (z_v,f^2) curve curls up, as happens near a critical frequency, an increase in the height of the equivalent triangular path has relatively less effect than in the case of a linear (z_v,f^2) curve. This means a relative increase in the wave frequencies calculated for points on a curve near a critical frequency. Thus it may be expected that the maximum usable frequencies calculated for an actual distribution of ionization density will be, for $a=1/3$, equal to or less than those calculated for $a=0$ only for very short distances. Elsewhere they will be greater, and may become as much as 20 percent or so greater, at the greater distances, than in the case where $a=0$. The calculation must in general be made for each individual distribution of ionization density.

The calculation can be made for a general distribution of ionization by assuming the (z_v,f^2) curve to be made up of a number, m, of linear intervals. In the nth such interval the ionization density, represented by f_0^2, is assumed to vary between f_n^2 and f_{n+1}^2, while the height varies from 0 to z_n. In this interval, then

$$z(f_0^2)=\frac{f_0^2-f_{n-1}^2}{f_n^2-f_{n-1}^2}\,z_n.$$

Then, if $x=f_0^2/f^2$ and $f_m^2=f^2$,

$$\frac{dz}{dx}=\frac{f^2 z_n}{f_{n+1}^2-f_n^2}$$

and is constant over the interval, so that eq 7 becomes

$$z_v=f^2\sum_{n=1}^{m}\frac{z_n}{f_n^2-f_{n-1}^2}\int_{\frac{f_{n-1}^2}{f^2}}^{\frac{f_n^2}{f^2}}\frac{1-a\dfrac{x^2}{(1+ax)^2}}{\sqrt{1-\dfrac{x}{1+ax}}}\,dx, \tag{11}$$

and eq 8 becomes

$$z_v'=f^2\sum_{n=1}^{m}\frac{z_n}{f_n^2-f_{n-1}^2}\int_{\frac{f_{n-1}^2}{f^2}}^{\frac{f_n^2}{f^2}}\sqrt{\frac{1+a\,\dfrac{(1-a)\cos^2\phi_1}{1-a\cos^2\phi_1}x}{1-(1-a)x}}\,dx. \tag{12}$$

Each of the integrals in the summation represents a part of the complete integral from 0 to $1/(1-a)$ and may be readily evaluated for a given a and cos ϕ_1. Both eq 11 and 12 are linear in z_n and may easily be evaluated. If we know only the (z_v,f) curve, eq 11 may be turned inside out and the known values of z_v used to form n equations in the n unknown values of z_n, which can then be substituted in eq 12 to obtain z_v'.

Calculations of the maximum usable frequencies for E- and F_2-layer transmission were made, using this method, for $a=0$ and $a=1/3$, for typical distributions of ionization density. The ratios of the maximum usable frequency for $a=1/3$ to that for $a=0$ are plotted, for various values of distance D, in figure 2. The solid graph is for F_2-layer and the dotted for E-layer transmission.

For F_2-layer transmission, it is seen that the maximum usable frequencies are not much greater for $a=1/3$ than for $a=0$ until the distance of transmission exceeds about 500 kilometers. For greater

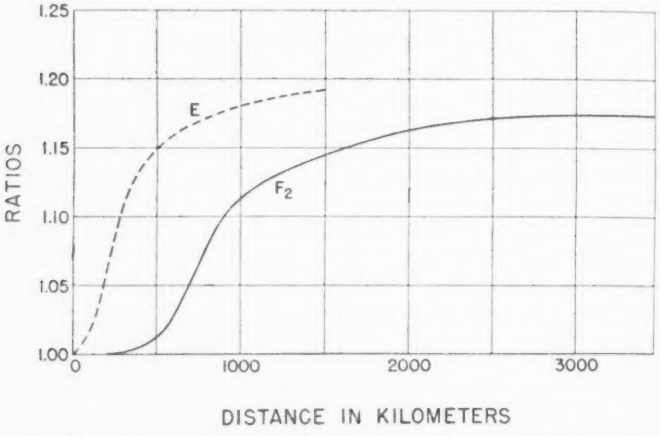


FIGURE 2.—*Ratios of maximum usable frequencies for $a=1/3$ to those for $a=0$, for E- and F_2-layer transmission.*

distances the maximum usable frequencies for $a=1/3$ are higher than those for $a=0$, becoming about 17 percent higher at the greater distances.

For E-layer transmission the ratio of maximum usable frequency for $a=1/3$ to that for $a=0$ rises much more rapidly with distance than in the case of F_2-layer transmission, and attains a somewhat greater value.

III. EFFECT OF THE EARTH'S MAGNETIC FIELD

The calculations so far have not included the effect of the earth's magnetic field. Since this field may affect considerably some high-frequency transmissions, it is necessary to consider its effect. Accordingly calculations were made for the transmission path and frequency, for which experimental results are given below.

The transmission path was practically perpendicular to the magnetic meridian, and was about 650 km long. The earth's field was assumed to have an intensity of 0.5 gauss, and a dip of 70°. For a frequency of 6.06 Mc/s, and a vertical-incidence virtual height of 130 km at the equivalent vertical-incidence frequency, table 1 gives the calculated oblique-incidence data. These were calculated on the basis of an average distribution of ionization density in the E region, using the magneto-ionic equations as published by Booker [9].

TABLE 1.—*Calculated oblique-incidence data*

Earth's field	Value of a	F_o at level of reflection	Angle of incidence	Equivalent vertical-incidence frequency	Vertical-incidence critical frequency
Gauss		*Mc/s*	*Degrees*	*Mc/s*	*Mc/s*
0	0	2.25	68.2	2.25	2.45
0	1/3	2.34	67.8	1.91	2.08
0.5	0	2.20	68.7	2.20	2.40
.5	1/3	2.31	68.1	1.89	2.06

The table gives, for $a=0$ and 1/3 and for $H=0$ and 0.5 gauss, the ionization density at the level of reflection, the angle of incidence of the waves which traverse the given path, the equivalent vertical-incidence frequency, and the vertical-incidence critical frequency which corresponds to a maximum usable frequency of 6.06 Mc/s over the given path, via E-layer transmission. It may be seen that the inclusion of the earth's field where $a=0$ makes a difference of but 2.0 percent, and where $a=1/3$ a difference of but 1.0 percent. The inclusion of the polarization term, however, makes a difference of 15.1 percent where $H=0$ and 14.2 percent where $H=0.5$ gauss.

Thus it should prove easy to decide between the Lorentz and Sellmeyer theories by measuring the vertical-incidence E-layer critical frequency at the times when the oblique-incidence E-layer transmission begins and fails. The experimental error should be considerably less than the difference between the two theories, either for calculations based on $H=0$ or for closer calculations based on $H=0.5$ gauss. The calculations for east-west transmission may be made for $H=0$ with an error much less than the probable experimental error.

A recent analysis has been published by Ratcliffe [10] for a parabolic distribution of ionization density in the ionosphere. The curves and discussion he gave for oblique-incidence radio transmission indicated little difference between the results for $a=0$ and $a=1/3$, especially for long-distance transmission via the F layer. Indeed, he stated that the difference would be greatest for a transmission distance of about 500 km.

The curves of figure 2 make this statement appear in error. It seems likely that Ratcliffe was misled by making calculations only for ranges below about 700 km, as published in his paper, for it is obvious that the greater separation between the two cases comes for considerably greater distances. It is possible, moreover, that the assumption of a parabolic distribution might also tend to lessen the difference at the shorter distances.

The results of the step-by-step analysis of typical E- and F-layer curves indicate the nature of the difference in maximum usable frequency according to the two theories.

IV. EXPERIMENTAL OBSERVATIONS

There is thus a clear distinction between the Lorentz theory ($a=1/3$) and the Sellmeyer theory ($a=0$) in the case of radio waves incident obliquely upon the ionosphere. It should therefore be possible to decide experimentally between the two theories.

There are available considerable published data on oblique-incidence radio transmission [11, 12, 13, 14, 15]. Experiments involving the E layer, which, as shown by figure 2, should provide conclusive results for relatively short transmission distances, are for the most part unreliable, usually because of the presence of sporadic-E reflections above the true E-layer critical frequency. Such an experiment was attempted by Farmer and Ratcliffe [11], with somewhat inconclusive results. Experiments involving the F_2 layer have for the most part been conducted over distances of less than 600 kilometers. They are subject to the difficulty that ionosphere conditions at the point of reflection are much more uncertain for the F layer than for the E layer. Also, the difference between the two theories would produce a discrepancy of but a few percent, which is of the same order of magnitude as the uncertainty of ionosphere conditions at the point of reflection. Much of the published data on radio transmission, therefore, is of little use in deciding whether $a=0$ or 1/3.

Since September 1935, the National Bureau of Standards has made continuous automatic field-intensity records of numerous radio stations. Results are available for the whole period from 1935 to 1940 for high-frequency international broadcasting station W8XAL (later WLWO), 6.06 Mc/s, at Mason, Ohio, 645 kilometers distant. Some results of these recordings were presented in a paper before the joint meeting of the International Scientific Radio Union and the Institute of Radio Engineers at Washington in April 1936, and other results have been used by investigators at the Bureau in studies of the ionosphere [16, 17].

The transmissions from W8XAL to Washington were propagated normally by way of the E layer during most of the daylight hours, and by way of the F layer at other times. The transitions from F- to E-layer transmission in the morning, and vice versa in the afternoon, were often well marked, since the vertical radiation pattern of the transmitting antenna was such as to provide substantially greater radiation for single-hop E- than for F-layer transmission. The transitions have been identified on a great many records by the following facts. (1) The regularity of occurrence of the change in intensity and character of fading on the records corresponds to the regularity of the diurnal and seasonal variations of the E-layer critical frequency. (2) There was a complete absence of any regular sudden change in local ionosphere characteristics (critical frequency, virtual height, or absorption), except for the E-layer critical frequency, at the time of the change on the record. (3) The beginning of propagation was observed when no other ionosphere layer but the E layer had sufficient ionization density to support transmission (as during the mornings of ionosphere-storm days). (4) The change on the record nearly always occurred about 15 minutes after the local E-layer critical frequency rose or fell through about 2.45 Mc/s. This time difference is the difference between the local times at Washington and at the midpoint of the path from W8XAL to Washington.

Figure 3 shows four examples of the change of layer on the W8XAL records. The record of May 12, 1938, was made when a severe ionosphere storm had considerably depressed the critical frequencies of the F_1 and F_2 layers. At the time when good transmission began, none of the layers except the E layer had even nearly enough ionization to support transmission. Records of February 21 and 22, 1938, were

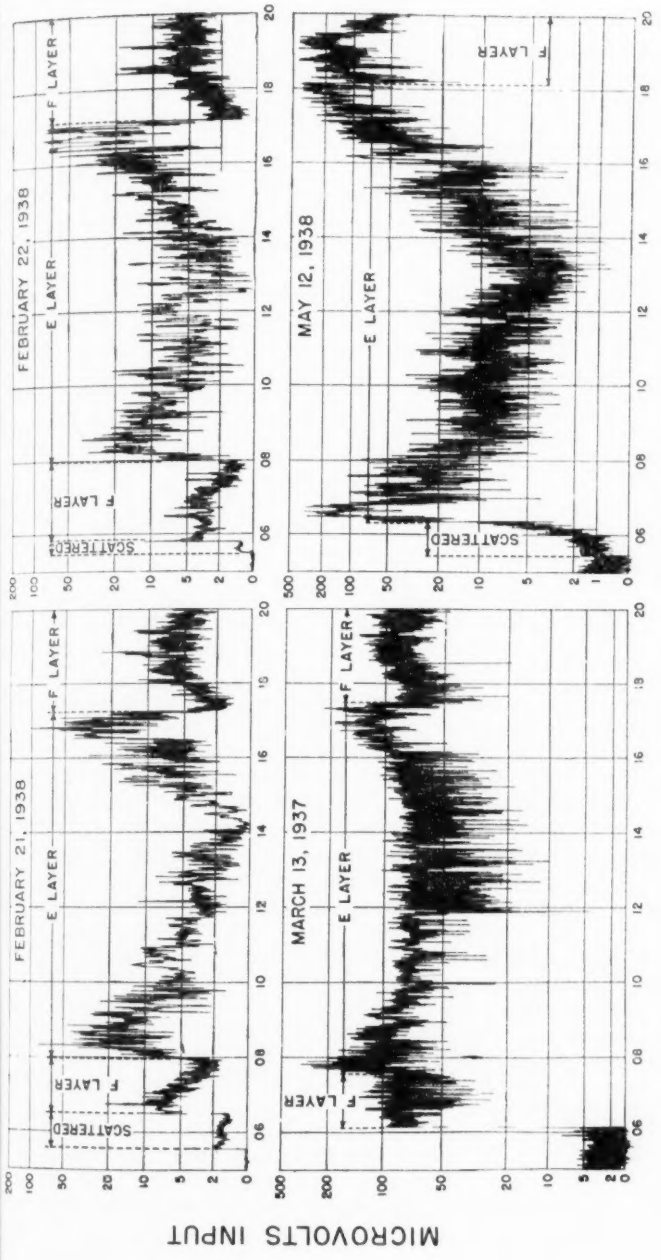


FIGURE 3.—*Transitions between E- and F_2-layer transmission from Mason, Ohio, to Washington, D. C.*

Station W8XAL, 6.06 Mc/s, 650 km distant.

made during a period when the night and morning F-layer transmissions from W8XAL were greatly absorbed. The transitions here are seen to be quite sharply defined and symmetrical about local noon at the midpoint of the transmission path. The record of March 13, 1937, is typical of most of the normal W8XAL records on which the transition can be seen. The transitions are not quite so sharp, but are easily identifiable, and the times may be measured within a few minutes. In the files of the National Bureau of Standards are 11 records similar to that of May 12, 1938, 13 similar to those of February 21 and 22, 1938, and several hundred showing the transition like that of March 13, 1937.

For the average regular distribution of ionization in the E layer, a maximum usable frequency of 6.06 Mc/s over a distance of 645 km corresponds to a vertical incidence $f_E°$ of about 2.40 Mc/s for $a=0$, and 2.10 Mc/s for $a=1/3$. The midpoint of the great-circle path between W8XAL and Washington is almost due west of Washington and the distance is short enough so that the $f_E°$ may be considered a function of local time only. Therefore the W8XAL transmissions are well adapted for deciding the value of a.

Since the E layer at the time of these observations is known to be chiefly electronic in character, it would seem that the presence of heavy ions cannot be invoked to explain a value of a different from 1/3 or 0. Also, it should be noted that it takes about 32 minutes for the $f_E°$ to rise in the morning, or fall in the evening, between 2.06 and 2.40 Mc/s. Also, the E^z reflections were so absorbed at these times, on the records analyzed, that there can be no question of confusing $E°$- with E^z-layer transmission.

Table 2 shows the result of comparison of 13 records, in each of which at least one transition was as well-defined as in the record of February 21 or 22 of figure 3. The values of a corresponding to the observed ratio of maximum usable frequency to vertical-incidence critical frequency are tabulated. In calculating these values of a, an average distribution of E-layer ionization was used instead of the actual distribution for each day, which varied somewhat from day to day. This may account for some of the variations in a.

The average value of a deduced from the 19 transitions listed in table 2 is -0.02, or $-1/50$, as against the theoretical value of 1/3. This value of a is not believed to be significantly different from zero.

A number of ionosphere storm days were studied, when no layer other than the E layer was sufficiently ionized, in the morning, to produce transmission. The transitions on many of these days were obscured by sporadic-E reflections. The mean value of a calculated on 7 days when the transition could be easily identified was -0.005, again not significantly different from zero.

Because of the great number of the records showing well-marked transitions which were not quite so clear as those just mentioned, the transitions on these records were studied statistically, rather than individually. The monthly averages of the morning and evening transition times were obtained, and the monthly average $f_E°$ at this time was determined from the vertical-incidence ionosphere data. Table 3 shows the average values of $f_E°$ at the midpoint of the transmission path, at the times when E-layer transmission began and failed, and the corresponding values of a. The mean value of a, deduced from 435 transitions studied in this group, was -0.06,

again not significantly different from zero. The negative value of a probably means that the height of the E layer was slightly greater than that assumed as an average.

TABLE 2.—*Values of "a" calculated from value of local f_E at time of change of layer for W8XAL, 6.06 Mc/s, 650 km distant*

Date (February 1938)	Local f_E in Mc/s		Calculated value of a	
	A. M.	P. M.	A. M.	P. M.
6	------	2.47	------	−0.08
14	2.33	2.43	+0.10	−.03
15	2.48	------	−.10	------
17	------	2.46	------	−.07
18	2.35	------	+.08	------
21	2.44	2.45	−.05	−.06
22	2.46	2.48	−.07	−.10
23	2.48	------	−.10	------
24	2.48	------	−.10	------
25	2.54	2.37	−.19	+.07
26	2.39	2.29	+.02	+.15
27	2.36	2.35	+.06	+.08
28	2.44	------	−.05	------

TABLE 3.—*Mean values of "a" calculated from values of local f_E at time of change of layer for W8XAL, 6.06 Mc/s, 650 km distant*

Month	Morning transition			Evening transition		
	Time	f_E°	a	Time	f_E°	a
1936						
May	------	------	------	1823	2.50	−0.13
June	------	------	------	1838	2.45	−.06
July	------	------	------	1830	2.42	−.02
September	0703	2.44	−0.05	1737	2.40	+.01
October	0738	2.52	−.16	1651	2.42	−.02
November	0810	2.40	+.01	1610	2.47	−.08
December	0831	2.48	−.10	1559	2.53	−.17
1937						
February	0805	2.45	−.06	1709	2.42	−.02
March	0726	2.47	−.08	1739	2.40	+.01
April	0637	2.40	+.01	1759	2.50	−.13
May	0604	2.42	−.02	1818	2.48	−.10
June	------	------	------	1849	2.40	+.01
July	------	------	------	1852	2.44	−.05
Mean a	------	------	−0.06	------	------	−0.06

V. CONCLUSIONS

For the cases considered in this paper, the data indicate that the Lorentz polarization term should not be included in the theory of oblique-incidence radio transmission by way of the ionosphere. Further experiments in different latitudes and with careful determinations of oblique and vertical-incidence penetration frequencies should be performed before this conclusion can be positively confirmed.

It is suggested that experiments of this type, involving wide ranges of frequencies and reflection heights, are less conducive to misleading results and ambiguous interpretations than are vertical-incidence experiments in the neighborhood of one frequency, which involve somewhat uncertain factors, such as the magnitude and direction of the earth's magnetic field at definite heights, and assumptions as to the physical constitution of the ionosphere.

VI. REFERENCES

[1] H. A. Lorentz, Theory of Electrons (G. E. Stechert, New York, N. Y., 1909).
[2] C. G. Darwin, Proc. Roy. Soc. (London), [A] **146,** 17 (1934).
[3] H. G. Booker and L. V. Berkner, *Constitution of the ionosphere and the Lorentz polarization term,* Nature **141,** 562 (1938).
[4] D. F. Martyn and G. H. Munro, *The Lorentz polarization term and the earth's magnetic field in the ionosphere,* Nature **141,** 159 (1938).
[5] E. V. Appleton, F. T. Farmer, and J. A. Ratcliffe, *Magnetic double refraction of medium radio waves in the ionosphere,* Nature **141,** 409 (1938).
[6] N. Smith, *Extension of normal-incidence ionosphere measurements to oblique-incidence radio transmission,* J. Research NBS **19,** 89 (1937) RP1013.
[7] N. Smith, *Application of vertical-incidence ionosphere measurements to oblique-incidence radio transmission,* J. Research NBS **20,** 683 (1938) RP1100.
[8] D. R. Hartree, Proc. Cambridge Phil. Soc. **27,** 143 (1930–31).
[9] H. G. Booker, *The propagation of wave packets incident obliquely upon a stratified doubly refracting ionosphere,* Roy. Soc. Phil. Trans. (London) **237A,** 411 (1938).
[10] J. A. Ratcliffe, *The effect of the Lorentz polarization term in ionospheric calculations,* Proc. Phys. Soc. (London) **51,** 747 (1939).
[11] F. T. Farmer and J. A. Ratcliffe, *Wireless waves reflected from the ionosphere at oblique incidence,* Proc. Phys. Soc. (London) **48,** 839 (1936).
[12] F. T. Farmer, C. B. Childs, and A. Cowie, *Critical frequency measurements of wireless waves reflected obliquely from the ionosphere,* Proc. Phys. Soc. (London) **50,** 767 (1938).
[13] E. V. Appleton and J. A. Ratcliffe, Proc. Roy. Soc. (London) **128,** 133 (1930).
[14] D. F. Martyn, R. O. Cherry, and A. L. Green, Proc. Phys. Soc. (London) **47,** 125 (1934).
[15] W. Crone, K. Kruger, G. Goubau, and J. Zenneck, *Echo measurements over long distances,* Hochfrequenztechnik Elektroakustik **48,** 1 (1936).
[16] J. H. Dellinger, *Sudden disturbances of the ionosphere,* J. Research NBS **19,** 111 (1937) RP1016.
[17] N. Smith and S. S. Kirby, *Critical frequencies of low ionosphere layers,* Phys. Rev. **51,** 890 (1937).

WASHINGTON, October 28, 1940.

U. S. DEPARTMENT OF COMMERCE NATIONAL BUREAU OF STANDARDS

RESEARCH PAPER RP1364

Part of Journal of Research of the National Bureau of Standards, Volume 26, February 1941

RAPID METHOD FOR DETERMINING ASCORBIC ACID CONCENTRATION

By Myron A. Elliott,[1] Alfred L. Sklar,[1] and S. F. Acree

ABSTRACT

In the course of investigations on the effect of citrus fruits upon the human organism it was found necessary to develop a method of determining the vitamin C content of the blood of a large number of individuals in a short time. In this method the plasma is not deproteinized but is diluted with 5-percent acetic acid and used directly in a specially constructed photoelectric comparator with the dye 2,6-dichlorophenolindophenol.[2] The ascorbic acid content of the plasma sample is read directly from empirical calibration curves determined by adding known amounts of ascorbic acid to blood plasma. The comparator is operated so as to compensate automatically for the usual variation in the turbidity and color of plasma samples. While the method is quite rapid, once it is in operation, some time must be spent in preparing calibration curves. The maximum total error in the determination on plasma samples containing from 0 to 3.5 mg of reduced ascorbic acid per 100 ml of plasma was estimated to be ± 0.1 mg/100 ml.

CONTENTS

	Page
I. Introduction	117
II. Apparatus	119
III. Procedure	121
IV. Calibration curves	122
1. Blood analyses	122
2. Urine analyses	125
V. Sources of error	126
VI. References	128

I. INTRODUCTION

In the course of investigations upon the effect of citrus fruits on the human organism it often becomes necessary to determine the ascorbic acid content in the blood of a large number of subjects in a relatively short time. Only by investigating a large group of test cases and another large group of controls can the effect of random statistical fluctuations and individual idiosyncrasies be made small. The time is limited by the necessity of performing the determinations on the entire group of subjects within the same fasting period and by the instability of ascorbic acid [3] in the presence of atmospheric oxygen [1].[4]

A chemical rather than a biological method must necessarily be used when speed of analysis is desired. The most specific and generally applicable reagent for ascorbic acid in use at present is the dye 2,6-dichlorophenolindophenol [2].[5] Tillmans [3] showed that the reducing substance in food that could be determined by titration with 2,6-

[1] Research Associates at the National Bureau of Standards, representing the Florida Citrus Commission.
[2] The sodium salt; also called sodium 2,6-dichlorobenzenoneindophenol.
[3] In this paper the words "ascorbic acid" always indicate the unoxidized compound.
[4] Figures in brackets indicate literature references at the end of this paper.
[5] See footnote 2.

dichlorophenolindophenol was approximately equal to the reported vitamin C content of the food. The identity of vitamin C with ascorbic acid was established by Waugh and King [4]. Tillmans and his associates [5], Harris and his associates [6] and Bessey and King [7] showed a close correlation to exist between the reducing power of plant and animal tissues on 2,6-dichlorophenolindophenol and the biological assay of the vitamin C content of these tissues. King [8] has given additional evidence to support the view that a properly conducted titration of natural products with the dye 2,6-dichlorophenolindophenol gives a reasonably specific test for vitamin C. This dye has been used regularly to titrate ascorbic acid in blood plasma [9, 10].

A simplification of existing methods for determining ascorbic acid in blood plasma by working directly upon the plasma without previously subjecting it to a deproteinization [9, 10] is introduced in the present paper. The time-consuming procedure of deproteinizing the plasma is open to the objection that some of the ascorbic acid is occluded by the precipitated proteins and is furthermore subject to atmospheric oxidation during the manipulations [11]. The primary purpose of the deproteinization is to obtain a clear solution for titrating or for absorption measurements in the photoelectric comparator. However, it has been shown [12] that turbid and colored solutions can be used, in general, with a photoelectric comparator for determinations of ascorbic acid with 2,6-dichlorophenolindophenol. It is shown in the present paper that when proper calibration curves have been made it is possible to determine directly the ascorbic acid content of blood plasma. Such analyses of the plasma are complicated, however, in about 5 percent of the samples by exceptional turbidity, which renders the determination less accurate. These samples may be analyzed by preparing special calibration curves, as discussed later, or by one of the longer chemical precipitation and extraction methods [9] if an accurate determination is necessary. Since the approximate determinations showed the ascorbic acid content to be uncorrelated with exceptional turbidity, these cases can merely be omitted in a study involving large numbers of analyses. It was observed that the number of plasma samples having excess turbidity was greater during the winter months, at which time the experiment was started, than later during the summer. From 5 to 10 percent of the samples were excessively turbid during February and March but less than 1 percent in May and June.

The amount of turbidity in blood samples is lowered if they are taken as long a time as possible after meals and the diet of the subjects is not too high in fats. No difficulties with hemolysis were experienced when the proper technique was used for obtaining the blood samples. If the samples are taken with a syringe, it is advisable to remove the needle from the syringe before the sample is transferred to a centrifuge tube in order to avoid rupturing the red blood cells by forcing them too rapidly through the needle. If the blood sample is stirred or shaken, these operations should be very gentle. The tube in which the sample is collected should be dry, except for 2 drops of a 20-percent solution of potassium oxalate.

II. APPARATUS

The photoelectric comparator was constructed in this laboratory along lines generally in use and is similar to one described by Evelyn [13]. It differs only in the light source, filters used, and control of the galvanometer deflection. The light source (*A*, fig. 1) was a 120-volt, 60-watt, clear, daylight, straight-filament, Mazda lamp (A-19, *CC*-b). The power was supplied through a constant-voltage regulator [6] (*R*) which kept the fluctuations of intensity of the light source to less than 1 percent when used on the regular a–c supply line. The image of the filament was focused through a condenser lens (*B*), through the filter (*F*), and the absorption cell (*C*) on the photoelectric cell (*D*). The absorption cells were carefully selected

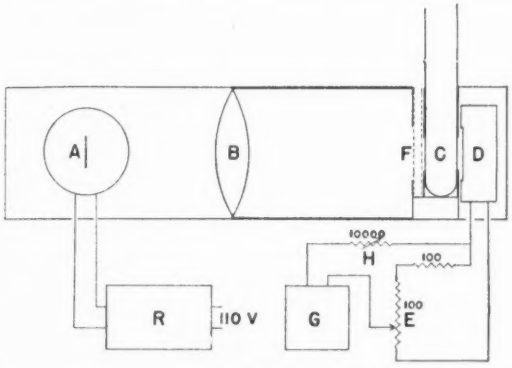


FIGURE 1.—*Schematic diagram of photoelectric comparator used for ascorbic acid analyses.*

calipered test tubes of 1-inch outside diameter by 6 inches long. When filled with a standard colored solution and rotated to a calibrated position marked on each test tube, the transmissions differed by less than 1 percent. This calibration took care of minor variations in wall thickness and roundness of the tubes. The photoelectric element was a Weston barrier-layer photronic cell, model 594, type 1. The galvanometer (*G*) was a Rubicon "spotlight" galvanometer having a coil resistance of 1,100 ohms, sensitivity of 0.005μa/mm, and critical damping resistance of 10,000 ohms. In figure 1, (*E*) is a 100-ohm slide-wire resistance by means of which the sensitivity of the galvanometer can be controlled without appreciably affecting the external resistance in series with the photronic cell. This method of controlling the deflection of the galvanometer has an advantage over the method which regulates the temperature of the filament of the light source [13] because, unlike the latter, it does not change the spectral distribution of the energy transmitted by the filters. In figure 1, (*H*) is a resistance used to effect critical damping of the galvanometer.

The light from the source passed through a yellow Wratten filter (No. 12) and then through a medium-blue glass filter before entering the absorption cell. The filter system, consisting of the two filters and the blue bulb, passed a spectral band with a transmission peak at

[6] Raytheon, 120-watt voltage regulator.

540 mμ and a width of 85 mμ at half peak. The band cut off sharply on the short-wave side, but tailed off on the long-wave side. The dye 2,6-dichlorophenolindophenol has in the acidified solution used in this work, a maximum absorption at 520 mμ and a half-peak width of 140 mμ, so that the light passed by the filters falls well within the region that the dye absorbs sensitively. The fraction of incident energy absorbed by a 1-in. layer of the blood plasma itself, in the dilution used, is about 10 percent at 600 mμ and increases to 25 percent at 515 mμ and 50 percent at 470 mμ. The filters were, accordingly, chosen to pass a spectral band toward the long-wave side of the region absorbed strongly by the unreduced dye (at pH 3) in order to have a minimum of absorption by the plasma solution alone.

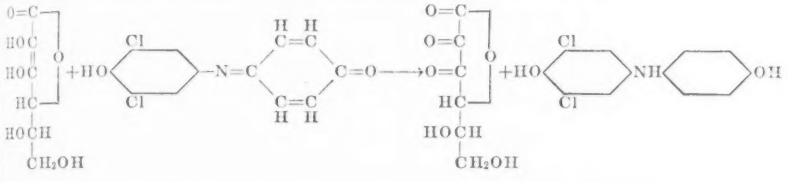
Filter systems transmitting more nearly homogeneous energy than that just described may be used with this comparator if some changes are made in the instrument. In some later work, the clear daylight lamp used as the light source was replaced by a colorless clear-bulb lamp of the same wattage and type, and a flat mirror was placed close behind the lamp so as to reflect a second image of the filament through the condensing lens. The 100-ohm slide-wire resistor (*E*, fig. 1) and the 100-ohm resistor, both in series with the photronic cell, were replaced by a 1,000-ohm slide-wire resistor with a vernier adjustment. The filter system consisted of the Corning glass filters No. 429 (medium-shade, blue-green) and No. 352 (H. R. Noviol) before the absorption cell and Wratten gelatin filter No. 62 (mercury green) between the absorption cell and the photronic cell. This filter combination gives a transmission peak at 530 mμ and a width of 28 mμ at half peak. An all-glass filter combination suggested by Corning Glass Works and consisting of their filters Nos. 352, 430, and 502 may also be used. This combination has a transmission peak at 518 mμ and a half-peak width of about 32 mμ. The substitution of the 1,000-ohm resistance for the 200-ohm resistance in series with the photronic cell will mean some sacrifice in the linearity of the response curve of the cell, even at the low radiant-flux densities used here. This situation can be improved, if necessary, by using a more sensitive galvanometer and leaving the resistance in series with the photronic cell at 200 ohms. One of the Rubicon high-sensitivity "spotlight" galvanometers could be used for this purpose. Their "spotlight" galvanometer, having a critical damping resistance of 40,000 ohms and a sensitivity of 0.0006 μa/mm, happened to be available and was found to be satisfactory. However, with the circuit shown in figure 1 their "spotlight" galvanometer having a critical damping resistance of 10,000 ohms and sensitivity of 0.0015 μa/mm would be better, both in sensitivity and period. The sensitivity of the photocell was improved by the use of one of the latest type barrier-layer photoelectric cells in the comparator.

The stock solution of 2,6-dichlorophenolindophenol was prepared by dissolving 100 mg of the dye crystals in warm distilled water, filtering, and diluting to 200 ml. The dye solution was kept in a refrigerator when not in use, and a fresh solution was prepared every few days. This was necessitated by the fact that after standing for 1 or 2 weeks (especially if warm) the dye is not quite completely bleached from reduction by ascorbic acid.

III. PROCEDURE

In determining the concentration of ascorbic acid, 10-ml samples of blood were collected in centrifuge tubes containing 2 drops of a 20-percent solution of potassium oxalate. After very gentle stirring, the blood was centrifuged for 10 minutes. Three (or, if necessary, two) milliliters of plasma was then transferred to a colorimeter cell and 15 (or 16) ml of 5-percent acetic acid was added, making a total volume of 18 ml. This gave a solution of pH 3.1. After stirring the solution, the cell was placed in the colorimeter and the slide wire (E), which was capable of being set to 1 part in 1,000, was carefully adjusted to give a deflection of 100 on the galvanometer. This procedure compensates for differences in the transmission of the original plasma samples, so that they may all be referred to the same calibration curve. The solution was then quickly poured into a test tube containing 1 ml of standard dye solution and at once poured back into the absorption cell. These mixing operations were performed in about 3 seconds. The galvanometer deflection can be read at any desired time intervals after 10 seconds from mixing, 20 seconds being adopted as the standard time necessary to allow the reaction to take place between the dye and the ascorbic acid in the plasma solution.

The chemical reaction is represented by the following equation:



Ascorbic acid. | 2,6-Dichlorophenolindophenol (blue in alkali, red in acid). | Dehydro-ascorbic acid. | Reduced indicator (colorless).

The determinations were made on groups of 20 samples. All the reagents were added in a uniform manner from transfer pipettes with enlarged tips and equipped with rubber bulbs. The necessary chemicals were added to a battery of 20 absorption cells while the first batch of blood samples was being centrifuged. A second batch of samples was centrifuged by the time the preceding batch was analyzed and the absorption tubes were again prepared. The operation was thus continuous and minimized the length of time during which the blood samples were kept. The analysis of a batch of 20 samples was completed within 40 minutes after they were taken from the ice box. The plasma samples originally were placed in the ice box within 10 minutes after they were collected, and the period of standing in the ice box was never more than 3 hours. The plasma was always removed from the centrifuge tube and acidified within 10 minutes after its separation from the whole blood. The stock dye solution was diluted to one-fourth (dye 4) its original strength when used for determinations on plasma with concentrations of ascorbic acid from 1 to 2 mg/100 ml, and to one-eighth (dye 8) of its original strength for plasma containing less than 1 mg/100 ml. The amount of ascorbic acid in the sample was read directly from the calibration curves in mg/100 ml of plasma.

IV. CALIBRATION CURVES

1. BLOOD ANALYSES

Figure 2 contains calibration curves prepared by various methods based on the use of 1 ml of dye 4 and 3-ml samples of the unknown solution. The data for curve 1 were obtained by adding successive increments of ascorbic acid to a 5-percent acetic acid solution containing no plasma and taking the galvanometer reading 20 seconds after mixing the solution with the dye. Curve 2 gives the results obtained by adding successive increments of ascorbic acid to the

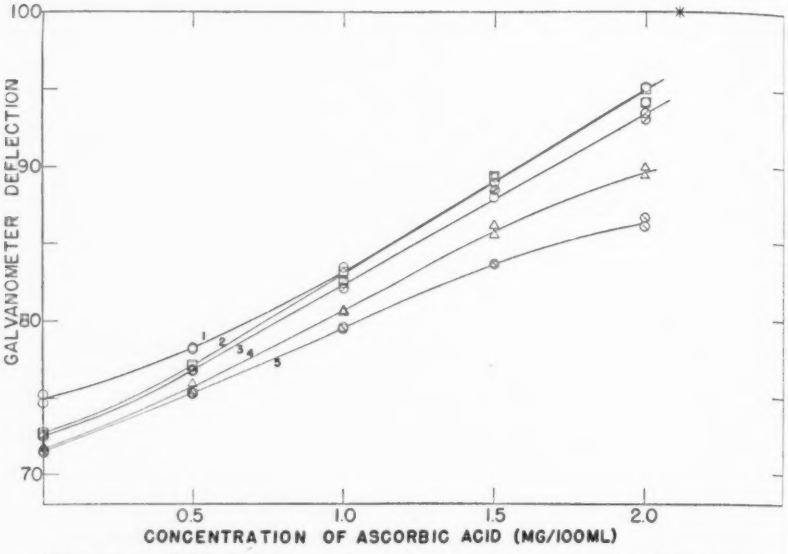


FIGURE 2.—*Calibration curves for photoelectric comparator, dye dilution 4.*

1. Solvent, 20-second reaction period. 2. Plasma, 30-second reaction period. 3. Plasma, 20-second reaction period. 4. Plasma, 15- to 30-second extrapolation. 5. Plasma, 10- to 20-second extrapolation.

solutions of acetic acid containing 3 ml of plasma free of ascorbic acid (see below) and allowing a 30-second interval between mixing with the dye and reading the galvanometer. Curve 3 is the same as 2, except that a 20-second interval was used instead of 30 seconds. Curve 4 differs from curve 2 in that the galvanometer readings were taken at 15 and 30 seconds and these values extrapolated linearly to zero time. Curve 5 was obtained similarly, except that 10- and 20-second readings were used for the extrapolation instead of the 15- and 30-second ones. The point designated by a star on the 100 ordinate represents the concentration of ascorbic acid which is stoichiometrically equivalent to 1 ml of the dye used. If the increments of ascorbic acid were totally oxidized within the time allowed for taking the readings, and if the plasma had no effect on the dye or on the ascorbic acid, then the calibration curves should all pass through the starred point. The fact that they do not shows that factors such as those just mentioned do complicate the reaction and points to the advantage of using an empirical calibration curve. For working with blood samples containing from 2 to 3.5 mg of ascorbic acid per 100 ml, a curve similar

to curve 3 but containing 2-ml plasma samples was prepared. For working with blood samples containing less than 1 mg of ascorbic acid per 100 ml, greater precision is obtained by using a more dilute dye solution. Figure 3 contains the calibration curve for 3-ml plasma samples, using the 20-second reaction period and dye dilution 8.

The preparation of a satisfactory blank plasma, for use in obtaining the calibration curves, presented some difficulties. The blank should contain no ascorbic acid but should have as nearly as possible the normal amount of the slowly reacting interfering substances. The

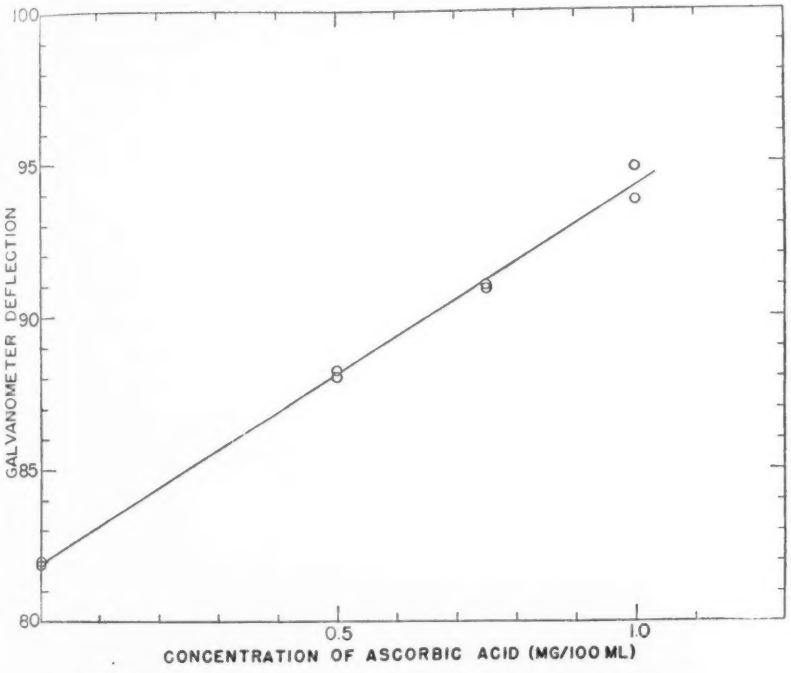


FIGURE 3.—*Calibration curve for blood plasma, dye dilution 8.*

blank plasma was prepared by two separate methods. The standard method was to titrate the fresh plasma, acidified with acetic acid, with dye as long as the photoelectric colorimeter showed the added dye to fade completely in 20 seconds. An alternative method was to let the plasma stand in the ice box for about 6 days and depend on the atmosphere to oxidize the ascorbic acid. These two independent methods for preparing plasma free of ascorbic acid gave samples so nearly the same that calibration curves prepared from them checked to within 0.1 mg of ascorbic acid per 100 ml of plasma.

Figure 2 shows that the calibration curves prepared from data which had been extrapolated to zero time (curves 4 and 5) deviated considerably from linearity in the region of higher concentration, while curves prepared from data obtained by using a straight 20- or 30-second reading (2 and 3) were nearly linear and came much closer to passing through the starred point. Important also is the

fact that the curves from the 10- to 20- and 15- to 30-second extrapolations have smaller slopes and hence give less sensitivity at high concentrations. The fact that the curves extrapolated to zero time are appreciably below those in which a definite reaction period was allowed shows that the reaction of ascorbic acid with the dye in blood plasma is not instantaneous. Figure 4 gives curves of galvanometer deflection versus time, for various solutions when 1 ml of dye 4 is added to them. These curves (and a duplicate check set not given here) furnish the data from which some of the points on the curves in figure 2 were obtained. The curves in figure 4 labeled 0, 1, and 2 were for 3-ml samples of blood plasma, diluted with 5-percent acetic acid in the regular manner, in which the original

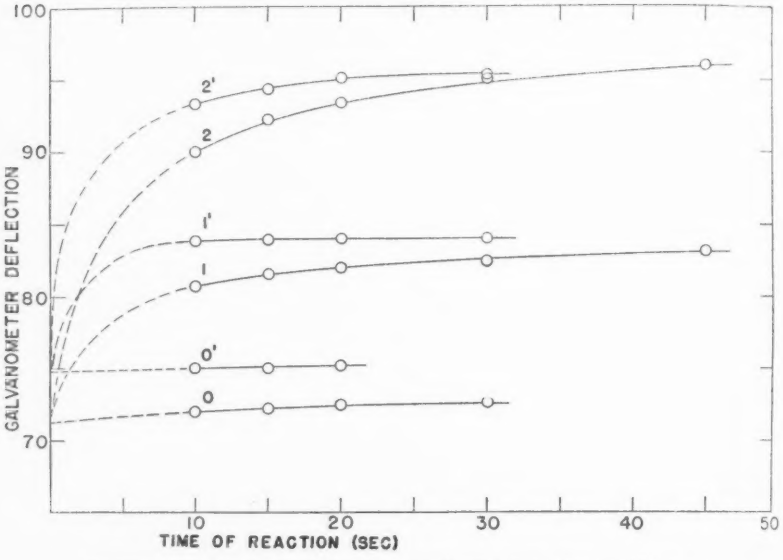


FIGURE 4.—*Rate of reaction of ascorbic acid with dye, dilution 4.*

Curves 0, 1, and 2 are for ascorbic acid concentrations of 0, 1, and 2 mg/100 ml, respectively, in acetic acid, blood plasma solution; curves 0′, 1′, and 2′ are the same as 0, 1, and 2, respectively, except no blood plasma was present.

ascorbic acid concentration was, respectively, 0, 1, and 2 mg/100 ml. The curves labeled 0′, 1′, and 2′ are given for comparison and were the same as 0, 1, and 2, except that they were made with the 5-percent acetic acid solution without the blood plasma present. It can be seen that at least 15 seconds is required for the reaction in the plasma solutions when the ascorbic acid concentration is of the order of 1 mg/100 ml, and at least 20 seconds is necessary when the concentration is 2 mg/100 ml. A comparison of the curves made with plasma solutions with those made with 5-percent acetic acid only show that the plasma reduces the rate of the reaction. As the ascorbic acid concentration approaches the stoichiometric equivalent of the dye, the reaction time lengthens appreciably. However, since the blood plasma also contains slowly reacting interfering substances [14], the reaction time must be kept to a minimum. A compromise must, accordingly, be reached to allow sufficient time for the ascorbic acid

to react but to minimize the reaction of the interfering substances. Calibration curves prepared on the basis of a 20-second reaction time were used in this work, since a linear curve was obtained for the type of plot used in figure 2 and 20 seconds appears to be the minimum time necessary for the oxidation of ascorbic acid in blood plasma. An extrapolation to zero time from readings taken on a 30- to 60-second basis gave essentially the same results as the 20-second readings. This extrapolation has an advantage in that it takes into account a second-order error due to slight differences in the amount of interfering substances between various blood plasma samples. Its chief disadvantage in this work is that it takes an appreciably longer time to analyze each sample. This time factor is important

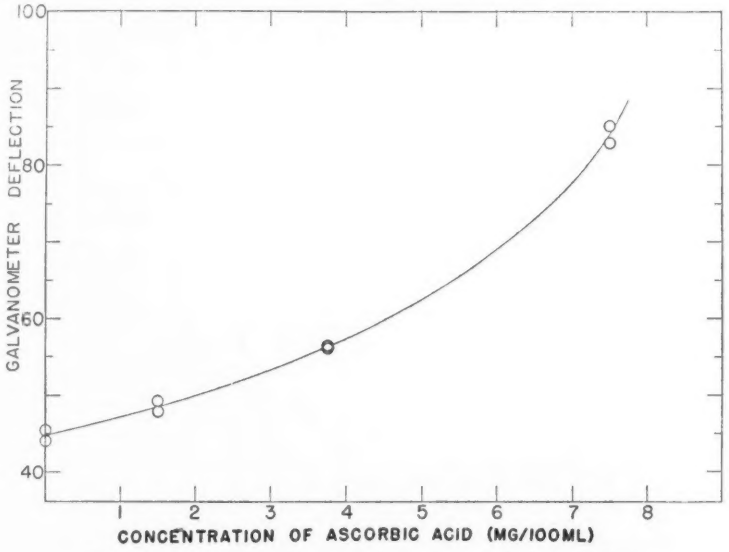


FIGURE 5.—*Calibration curve for urine, stock dye solution.*

when it is necessary to analyze large numbers of samples before oxidation by the air has a chance to affect appreciably their ascorbic acid content. With the method described here, two operators can analyze 40 blood samples in an hour.

2. URINE ANALYSES

The photoelectric colorimeter may also be used to determine the ascorbic acid in urine [12, 15], which contains many more interfering substances than does blood plasma. These substances are partially removed by precipitation with barium acetate [16, 17]. In the procedure for urine, ½ ml of glacial acetic acid and 2 g of dry barium acetate are added to 10 ml of the fresh sample. After shaking, the sample is centrifuged and 1 to 3 ml (depending on the ascorbic acid contained) is analyzed in the colorimeter in the same manner as the blood plasmas. However, instead of the 20-second reaction time used for blood analyses, it is necessary to extrapolate to zero time in order to minimize the effect of the slowly reacting interfering substances which remain even after the precipitation with barium

acetate. In these experiments a 10- to 20-second extrapolation was found most satisfactory. Dye dilution 4 is used for urine samples containing less than 2 mg of ascorbic acid per 100 ml of urine, and the undiluted dye is used for samples containing 2 mg or more per 100 ml. Here again empirical calibration curves were found to be preferable to the stoichiometric method. These curves were prepared by adding known increments of ascorbic acid to urine that contained no reduced ascorbic acid. Figure 5 contains the calibration curve prepared by using the undiluted dye and 3-ml samples of urine. The urine analyses may be made in batches of 20 and with about the same speed as the analyses of ascorbic acid in blood.

V. SOURCES OF ERROR

As mentioned previously, those plasma samples which showed excessive turbidity were discarded. They could, however, be analyzed if separate calibration curves were made. A small amount of turbidity causes no difficulty, since the transmission is always adjusted to the 100 mark before the addition of dye. Excessive turbidity, however, occasionally appears in some plasma samples. If a 1-inch layer of a 5-to-1 dilution of plasma transmits less than 65 percent as much light as a 1-inch layer of water, it is here considered to be excessively turbid. Calibration curves prepared on excessively turbid solutions are shifted upward into the region of higher transmission with respect to those formed on clear or slightly turbid samples, the shift being largest for low concentrations of ascorbic acid. The effect is probably not directly due to the excessive turbidity but rather to the absorption of dye by the colloidal particles causing the turbidity. An absorption phenomenon accounts for the experimental facts, which are in the opposite direction to what would be expected solely from the optical effect of a large amount of turbidity. When working with a turbid sample care must be taken to see that it stands long enough (5 to 15 minutes) after it is diluted with acetic acid for the turbidity to become constant with respect to time.

The effects described above might, of course, produce small errors even in slightly turbid plasmas. A good test for such errors, arising from variation in the natural color and in the turbidity, is obtained by comparing the final galvanometer readings for a series of plasmas, free of reduced ascorbic acid, to each of which 1 ml of dye 8 has been added. The results on plasmas of 11 individuals taken at random are given in table 1. The first column gives the original percentage transmission, compared with distilled water, of a 5-to-1 dilution of plasma with 5-percent acetic acid.

TABLE 1.—*Effect of turbidity on calibration curves*

$T_{orig.}/T_{H_2O}$	Final galvanometer reading	$T_{orig.}/T_{H_2O}$	Final galvanometer reading
Percent		*Percent*	
74	82.0	80	82.7
76	82.1	80	82.4
78	81.3	81	81.6
78	81.9	83	81.7
79	82.0	85	81.1
79	81.9		
			Mean....81.9±0.3

The galvanometer is then adjusted to 100 and 1 ml of of dye 8 added, the resulting galvanometer reading being recorded in the second column of the table. The latter galvanometer readings, except for the errors discussed above, should all be identical. The average deviation of 0.3 from the mean (81.9) corresponds (see fig. 3) to an uncertainty of approximately 0.03 mg/100 ml of ascorbic acid in the undiluted plasma.

Atmospheric oxidation of the ascorbic acid will introduce systematic errors if not controlled. However, in these experiments the samples were never kept more than 3 hours in the refrigerator nor, upon dilution and immediate acidification, for as much as 40 minutes at room temperatures. It is known [1, 9] that no appreciable oxidation by air takes place in 3 hours in the refrigerator. Furthermore, no difference in the ascorbic acid content could be detected in diluted and acidified plasma samples because of standing for 40 minutes at room temperatures. The only other possibility for oxidation by air is in the centrifuge and while the plasma samples are being acidified, especially since ascorbic acid is known to be less stable in plasma than in whole blood. Although appreciable oxidation of the ascorbic acid in blood plasma is known to take place when the plasma stands for more than 4 hours at room temperature [1, 18], the amount of oxidation in the 15 minutes taken to centrifuge and acidify a batch of 20 samples will be small, especially since it takes a part of the 15 minutes for the samples to warm up to room temperature. Roe [19] has discussed various sources of error in his and other methods.

Errors due to a slow fading of the dye in 5-percent acetic acid solution, to nonlinearity of the photoelectric-cell response, and possible incompleteness of the reaction of ascorbic acid with dye in 20 seconds are practically eliminated by the use of empirical calibration curves instead of stoichiometric equivalents. Air bubbles introduced by mixing the solutions cause no difficulty in the transmission measurements if the readings are taken 10 seconds or more after mixing.

The maximum total error in these determinations of ascorbic acid in plasma containing 0 to 3.5 mg/100 ml is estimated as ± 0.1 mg/100 ml.

This work was made possible by a grant from the Florida Citrus Commission to the National Research Council and the University of Maryland. It was carried out under the direction of a research advisory committee consisting of Sterling V. Mead, Dental Clinic of Mead, Newton, Lynch, and Diggman, Washington, D. C.; Lawrence H. James, Department of Bacteriology, University of Maryland; S. F. Acree, National Bureau of Standards; Frank R. Klune, Hospital of the District of Columbia Reformatory, Lorton, Va.; the National Research Council, Washington, D. C.; and L. W. Marvin, chairman, Scientific Advisory Committee, Florida Citrus Commission, Lakeland, Fla.

We are indebted to J. P. Porch, Health Department of the District of Columbia, and to Oscar B. Hunter, physician, Washington, D. C., for supplying samples of serum and blood used in preliminary tests of this method; and to Joseph Rosin, of Merck Co., for supplying crystalline ascorbic acid and 2,6-dichlorophenolindophenol.

VI. REFERENCES

[1] H. Borsook, H. W. Davenport, C. E. P. Jeffreys, and R. C. Warner, *The oxidation of ascorbic acid and its reduction in vitro and in vivo.* J. Biol. Chem. **117**, 237 (1937).

[2] The Vitamins, A Symposium, chapter 20, 1939. (O. A. Bessey) Am. Med. Assn.

[3] J. Tillmans, P. Hirsch, and W. Hirsch, *The reducing property of plant foods and its relation to vitamin C,* Z. Untersuch. Lebensmitt. **63,** 1 (1932).

[4] W. A. Waugh and C. G. King, *The chemical nature of vitamin C,* Science **75,** 357 (1932); *The isolation and identification of vitamin C,* J. Biol. Chem. **97,** 325 (1932).

[5] J. Tillmans, P. Hirsch, and R. Vanbel, *The reducing value of plant tissues and its indication of vitamin C,* Z. Untersuch. Lebensmitt. **65,** 145 (1933).

[6] L. J. Harris and S. N. Ray, *Specificity of hexuronic (ascorbic) acid as anti-scorbutic factor,* Biochem. J. **27,** 580 (1933).

[7] O. A. Bessey and C. G. King, *The distribution of vitamin C in plant and animal tissues and its determination,* J. Biol. Chem. **103,** 687 (1933).

[8] C. G. King, *Vitamin C, ascorbic acid,* Physiol. Rev. **16,** 238 (1936).

[9] R. L. Mindlin and A. M. Butler, *The determination of ascorbic acid in plasma; a macromethod and micromethod.* J. Biol. Chem. **122,** 673 (1937–38).

[10] A. F. Abt, C. J. Farmer, and I. M. Epstein, *Normal cevitamic (asorbic acid) determinations in blood plasma and their relationship to capillary resistance,* J. Pediat. **8,** 1 (1936). Caroline C. Sherman and H. C. Sherman, *The vitamins,* in Annual Review of Biochemistry, Stanford University, California, Stanford University Press **6,** 346 (1937). C. J. Farmer and A. F. Abt, *Determination of reduced ascorbic acid in small amounts of blood,* Proc. Soc. Exptl. Biol. Med. **34,** 146 (1936). M. Pijoan and F. Klemperer, *Determination of blood ascorbic acid,* J. Clin. Investigation **16,** 443 (1937).

[11] R. Lemberg and J. W. Legge, *Determination of ascorbic acid in the presence of hemoglobin,* J. Proc. Roy. Soc. N. S. Wales **72,** 62 (1938).

[12] O. A. Bessey, *A method for the determination of small quantities of ascorbic acid and dehydroascorbic acid in turbid and colored solutions in the presence of other reducing substances,* J. Biol. Chem. **126,** 771 (1938).

[13] K. A. Evelyn, *A stabilized photoelectric colorimeter with light filters,* J. Biol. Chem. **115,** 63 (1936).

[14] A. Emmerie and M. van Eekelen, *The chemical determination of vitamin C with removal of interfering reducing and colored substances.* Biochem. J. **28,** 1153 (1934).

[15] K. A. Evelyn, H. T. Malloy, and C. Rosen, *The determination of ascorbic acid in urine with the photoelectric colorimeter,* J. Biol. Chem. **126,** 645 (1938).

[16] M. Van Eekelen, *Ascorbic acid and thiosulphate in urine.* Acta Brevia Neerland. Physiol. Pharmacol. Microbiol. **4,** 137 (1934). *Ascorbic acid and thiosulphate in urine,* Nature **135,** 37 (1935).

[17] M. Van Eekelen and M. Heinemann, *Critical remarks on the determination of urinary excretion of ascorbic acid,* J. Clin. Investigation **17,** 293 (1938).

[18] M. Cushman and A. M. Butler, *Use of cyanide in the determination of ascorbic acid,* Proc. Soc. Exptl. Biol. Med. **39,** 534 (1938).

[19] J. H. Roe and J. M. Hall, *The vitamin C content of urine,* J. Biol. Chem. **131,** 329 (1939).

WASHINGTON, December 6, 1940.

U. S. DEPARTMENT OF COMMERCE NATIONAL BUREAU OF STANDARDS

RESEARCH PAPER RP1365

Part of Journal of Research of the National Bureau of Standards, Volume 26, February 1941

BOILING POINTS OF BENZENE, 2,2,3-TRIMETHYLBUTANE, 3-ETHYLPENTANE, AND 2,2,4,4-TETRAMETHYLPENTANE WITHIN THE RANGE 100 TO 1,500 MILLIMETERS OF MERCURY

By Edgar Reynolds Smith

ABSTRACT

By the comparative method of Świętosławski, with water for the reference standard, data were obtained from which were developed the following equations to express the relationship between temperature and vapor pressure in the range of pressures from 100 to 1,500 mm Hg.

For benzene,

$$\log_{10}p=6.905216-\frac{1211.215}{220.870+t}.$$

For 2,2,3-trimethylbutane,

$$\log_{10}p=6.799682-\frac{1204.997}{226.615+t}.$$

For 3-ethylpentane,

$$\log_{10}p=6.873058-\frac{1249.825}{219.595+t}.$$

For 2,2,4,4-tetramethylpentane, no convenient single equation was found to fit the data with satisfactory accuracy over the entire range of pressures. By dividing the data into two ranges, satisfactory reproduction was obtained with the following equations:

(a) From 100 to 450 mm,

$$\log_{10}p=6.643408-\frac{1231.620}{204.975+t}.$$

(b) From 430 to 1,500 mm,

$$\log_{10}p=6.860684-\frac{1368.925}{221.679+t}.$$

In these equations, p is the vapor pressure, in standard millimeters of mercury, exerted by the substance at the temperature t in degrees centigrade.

The method, apparatus, and procedure used in this work have been described in a previous paper [5].[1]

From observations at various pressures of t_s, the boiling point of the substance under consideration, and t_w, the corresponding boiling point of water, the constants of equations of the type

$$t_s=a+bt_w+ct^2_w \tag{1}$$

were determined by the method of least squares. For 2,2,4,4-tetramethylpentane, however, the simple quadratic form of equation was insufficient and an additional cubic term was found necessary to

[1] Figures in brackets indicate the literature references at the end of this paper.

reproduce the data with adequate accuracy. Using the equations thus obtained, values of t_s corresponding to a series of reference values of t_w in the measured range were computed. The reference values of t_w and the corresponding reference pressures adopted for comparative ebulliometric measurements over the range of 100- to 2,000-mm pressure have been previously tabulated [5] and were taken from the compilation by Osborne, Stimson, and Ginnings [3].

The computed values of t_s at these reference pressures were used to obtain a relationship of the form

$$\log p = A - \frac{b}{C+t}, \tag{2}$$

which is explicit in either temperature or pressure. The symbol "log" is used to denote the logarithm to the base 10. Also

$$t_n - t = \frac{b(\log 760 - \log p)}{(A - \log 760)\ (A - \log p)}, \tag{3}$$

in which t_n is the normal boiling point and t is the boiling point at the pressure p. Equation 3 is useful for calculating the normal boiling point from a boiling point measured at any pressure within the range for which eq 2 is applicable.

The benzene used for the measurements was the fifth fraction from a fractional distillation described in a former paper [6]. The 2,2,3-trimethylbutane, 3-ethylpentane, and 2,2,4,4-tetramethylpentane were supplied by D. B. Brooks, of the Automotive Power Plants Section of this Bureau. Their preparation and properties have been described in other papers [1, 2].

Temperatures were measured to 0.001° C., and the equations expressing the data are given so that calculations may be made with them to that precision. The accuracy of the results, however, depends on the purity of the measured substances as well as on the method of measurement, calibrations of instruments, and experimental technique. From the methods of preparation and the observations as to the behavior of the substances with respect to distillation and freezing, which are given in the references cited, it is evident that all the substances were of a high degree of purity. However, no quantitative estimates of their purities are possible and no exact measure of the accuracy of the results can be made. It is believed that the results are accurate to better than 0.01°.

Benzene.—The measured corresponding boiling points of benzene and water are given in table 1. The expression for the boiling point of benzene as a function of the corresponding boiling point of water is

$$t = -26.8739 + 0.979949 t_w + 0.00089772 t_w^2. \tag{4}$$

The average deviation of the 16 measurements from eq 4 is 0.004° and the greatest deviation is 0.009°. The normal boiling point calculated from eq 4 is 80.098°. The difference between this and the value of 80.094° reported in an earlier paper [6] on measurements over the more limited range of 660 to 860 mm is just equal to the average deviation of the experimental values from those calculated by eq 4. This agreement is adequate. The temperatures calculated from eq 4 for the standard reference pressures are given in the second column

of table 5. The equation found to represent the pressure-temperature relationship is

$$\log p = 6.905216 - \frac{1211.215}{220.870 + t} \tag{5}$$

and reproduces the data given in the first and second columns of table 5, with an average deviation of 0.04 mm and a greatest deviation of 0.28 mm at the highest pressure. Values of the vapor pressure of benzene caculated by eq 5 for 30, 40, 50, 60, 70 and 80° C. yield an average deviation of 0.15 mm from the values tabulated by Scatchard, Wood, and Mochel [4] from their measurements on benzene over this range of temperatures. The equation

$$t_n = t + 300.968\frac{(2.880814 - \log p)}{(6.905216 - \log p)}$$

may be used to calculate the normal boiling point of a sample of benzene from the boiling point measured at any pressure in the range of 100 to 1,500 mm.

TABLE 1.—*Corresponding boiling points of benzene and water*

Boiling point		Boiling point	
Benzene	Water	Benzene	Water
°C	°C	°C	°C
33.248	58.243	70.604	91.758
37.707	62.344	75.484	96.010
42.977	67.158	79.774	99.723
48.440	72.087	85.483	104.630
53.015	76.201	90.224	108.678
57.677	80.366	95.339	113.015
60.877	83.200	100.493	117.358
65.276	87.084	105.052	121.168

2,2,3-Trimethylbutane (*triptane*).—The measured corresponding boiling points of 2,2,3-trimethylbutane and water are given in table 2. The expression found for the boiling point in terms of the corresponding boiling point of water is

$$t = -30.1869 + 1.004715t_w + 0.00105873t_w^2. \tag{6}$$

The average deviation of the 17 measurements from eq 6 is 0.003° and the greatest deviation is 0.008°. The normal boiling point calculated from eq 6 is 80.872°. The temperatures calculated by eq 6 to correspond with the standard reference pressures are given in the third column of table 5. The equation found to represent these pressure-temperature data is

$$\log p = 6.799682 - \frac{1204.997}{226.615 + t}, \tag{7}$$

with an average deviation of 0.07 mm and one exceptionally large deviation of 0.48 mm at the highest pressure. For calculating the normal boiling point from the boiling point measured at any pressure in the range of 100 to 1,500 mm, the equation,

$$t_n = t + 307.486\frac{(2.880814 - \log p)}{(6.799682 - \log p)}$$

may be used.

TABLE 2.—*Corresponding boiling points of 2,2,3-trimethylbutane and water*

Boiling point		Boiling point	
2,2,3-Trimethylbutane	Water	2,2,3-Trimethylbutane	Water
°C	°C	°C	°C
25.774	52.772	71.727	92.433
31.597	57.954	76.835	96.674
36.385	62.186	81.214	100.283
41.785	66.923	87.030	105.042
47.424	71.806	92.126	109.178
52.186	75.910	97.335	113.379
57.301	80.284	102.430	117.455
61.862	84.153	107.262	121.298
66.357	87.941		

3-Ethylpentane.—The comparative boiling-point data for 3-ethylpentane and water are given in table 3. The equation found for the boiling point of 3-ethylpentane in terms of the corresponding boiling point of water is

$$t=-18.2932+1.019200\ t_w+0.00098408\ t_w^2. \tag{8}$$

The average deviation of the 17 measurements from eq 8 is 0.002° and the greatest deviation is 0.005°. The normal boiling point calculated from eq 8 is 93.468°. The temperatures corresponding to the standard reference pressures are given in the fourth column of table 5. The equation representing the pressure-temperature relationship is

$$\log p=6.873058-\frac{1249.825}{219.595+t}, \tag{9}$$

which reproduces the data for 3-ethylpentane given in table 5 with an average deviation of 0.03 mm and a greatest deviation of 0.09 mm at the highest pressure. The normal boiling point of a sample of 3-ethylpentane may be obtained from a boiling point measured at any pressure between 100 and 1,500 mm by the equation

$$t_n=t+313.063\ \frac{(2.880814-\log p)}{(6.873058-\log p)}.$$

TABLE 3.—*Corresponding boiling points of 3-ethylpentane and water*

Boiling point		Boiling point	
3-Ethylpentane	Water	3-Ethylpentane	Water
°C	°C	°C	°C
37.486	52.103	86.992	94.652
43.732	57.652	91.730	98.571
49.337	62.576	94.846	101.134
54.572	67.141	100.618	105.851
60.567	72.324	104.967	109.384
65.592	76.635	110.175	113.589
70.493	80.806	114.899	117.379
75.785	85.282	120.047	121.486
81.541	90.113		

2,2,4,4-Tetramethylpentane.—The measured corresponding boiling points of this substance and water are given in table 4.

TABLE 4.—*Corresponding boiling points of 2, 2, 4, 4-tetramethylpentane and water*

Boiling point		Boiling point	
2,2,4,4-Tetramethylpentane	Water	2,2,4,4-Tetramethylpentane	Water
°C	°C	°C	°C
60.711	51.938	111.630	91.991
67.256	57.285	117.434	96.368
72.665	61.653	122.436	100.115
78.581	66.389	128.684	104.759
85.241	71.657	134.261	108.874
90.477	75.754	140.081	113.139
95.955	80.007	146.065	117.494
100.963	83.867	151.202	121.204
105.951	87.685		

The equation found to represent the corresponding data is

$$t=1.7064+1.049344t_w+0.00178137t^2_w-0.0000021712t^3_w, \tag{10}$$

which shows an average deviation of 0.004° and a greatest deviation of 0.009° from the 17 experimental values. The normal boiling point calculated from eq 10 is 122.283°. The temperatures calculated from eq 10 to correspond with the standard reference pressures are given in the fifth column of table 5. A single equation of the form of eq 2 would not adequately reproduce these pressure-temperature data over the entire range. By dividing the data into two ranges, satisfactory reproduction was obtained with the following two equations:

(a) From 100 to 450 mm,

$$\log p=6.643408-\frac{1231.620}{204.975+t'} \tag{11}$$

with average and greatest deviations of 0.02 and 0.04 mm, respectively.

(b) From 430 to 1,500 mm,

$$\log p=6.860684-\frac{1368.925}{221.679+t'} \tag{12}$$

with average and greatest deviations of 0.06 and 0.13 mm, respectively. The equation

$$t_n=t+343.962\frac{(2.880814-\log p)}{(6.860684-\log p)}$$

may be used to obtain the normal boiling point from a boiling point measured at any pressure within the range of 430 to 1,500 mm.

TABLE 5.—*Boiling points of benzene, 2,2,3-trimethylbutane, 3-ethylpentane, and 2,2,4,4-tetramethylpentane at standard reference pressures*

Pressure	Boiling point			
	Benzene	2,2,3-Trimethylbutane	3-Ethylpentane	2,2,4,4-Tetramethylpentane
mm	°C	°C	°C	°C
92.52	24.368	22.696	35.127	58.356
118.06	29.739	28.275	40.740	64.448
149.40	35.155	33.907	46.402	70.611
187.57	40.616	39.593	52.113	76.844
233.72	46.121	45.331	57.873	83.145
289.13	51.672	51.122	63.682	89.511
355.22	57.267	56.966	69.541	95.943
433.56	62.908	62.863	75.449	102.438
525.86	68.593	68.813	81.406	108.994
633.99	74.323	74.816	87.412	115.610
760.00	80.098	80.872	93.468	122.283
906.06	85.918	86.981	99.572	129.014
1,074.58	91.783	93.142	105.726	135.799
1,268.03	97.693	99.357	111.929	142.638
1,489.14	103.647	105.625	118.182	149.528

Values of the temperature and rates of change of pressure with temperature at even values of the pressure are given in table 6 for all four substances.

TABLE 6.—*Values of pressure, temperature, and rates of change of pressure with temperature for benzene, 2,2,3-trimethylbutane, 3-ethylpentane, and 2,2,4,4-tetramethylpentane*

Pressure	Benzene		2,2,3-Trimethylbutane		3-Ethylpentane		2,2,4,4-Tetramethylpentane	
	Temperature	*dp/dt*	Temperature	*dp/dt*	Temperature	*dp/dt*	Temperature	*dp/dt*
mm	°C	*mm*/°C	°C	*mm*/°C	°C	*mm*/°C	°C	*mm*/°C
50	• 11.78	• 2.58	• 9.63	• 2.49	• 21.96	• 2.47	• 44.12	• 2.29
100	26.054	4.57	24.443	4.40	36.882	4.38	60.266	4.03
200	42.198	8.06	41.242	7.73	53.768	7.70	78.653	7.05
300	52.660	11.18	52.154	10.71	64.719	10.68	90.641	9.74
400	60.601	14.08	60.452	13.47	73.036	13.44	99.780	12.21
500	67.086	16.82	67.236	16.07	79.830	16.05	107.254	14.57
600	72.611	19.43	73.021	18.54	85.620	18.54	113.634	16.82
700	77.450	21.94	78.094	20.92	90.692	20.92	119.224	18.99
760	80.098	23.40	80.872	22.30	93.468	22.32	122.283	20.25
800	81.773	24.36	82.629	23.21	95.225	23.23	124.219	21.08
900	85.691	26.71	86.742	25.43	99.334	25.46	128.749	23.10
1,000	89.283	28.99	90.516	27.59	103.102	27.64	132.902	25.07
1,200	95.702	33.39	97.265	31.74	109.837	31.82	140.327	28.86
1,400	101.340	37.61	103.200	35.71	115.755	35.83	146.851	32.49
1,600	• 106.39	• 41.67	• 108.52	• 39.53	• 121.06	• 39.68	• 152.696	• 35.98

• Extrapolated value.

REFERENCES

[1] D. B. Brooks, F. L. Howard, and H. C. Crafton, Jr., J. Research NBS **24,** 33 (1940) RP1271.
[2] F. L. Howard, J. Research NBS **24,** 677 (1940) RP1306.
[3] N. S. Osborne, H. F. Stimson, and D. C. Ginnings, J. Research NBS **23,** 261 (1939) RP1229.
[4] G. Scatchard, S. E. Wood, and J. M. Mochel, J. Am. Chem. Soc. **61,** 3206 (1939).
[5] E. R. Smith, J. Research NBS **24,** 229 (1940) RP1280.
[6] E. R. Smith and H. Matheson, J. Research NBS **20,** 641 (1938) RP1097.
[7] W. Świętosławski, Ebulliometry (Chemical Publishing Co. of New York, Inc., 1937).

WASHINGTON, November 25, 1940.

U. S. DEPARTMENT OF COMMERCE NATIONAL BUREAU OF STANDARDS

RESEARCH PAPER RP1366

Part of Journal of Research of the National Bureau of Standards, Volume 26, February 1941

INFLUENCE OF STRESS ON THE CORROSION PITTING OF ALUMINUM BRONZE AND MONEL METAL IN WATER

By Dunlap J. McAdam, Jr. and Glenn W. Geil

ABSTRACT

Cyclic stress tends to increase the size and sharpness of corrosion pits in aluminum bronze and monel metal and thus tends to increase the rate of lowering of the fatigue limit by corrosion. The form of corrosion pits in aluminum bronze is affected by the duplex microstructure. Pits in monel metal are not appreciably influenced by the microstructure. Curves of decrease of the fatigue limit, and constant-damage diagrams derived from these curves, are very different from those obtained with steels. These differences may be attributed to the fact that the rate of corrosion of aluminum bronze and monel metal, unlike that of steel, is anodically controlled. Steady stress tends to increase the rate of corrosion pitting of aluminum bronze but has little apparent effect on monel metal.

CONTENTS

	Page
I. Introduction	136
II. Materials and method of investigation	136
III. Lowering of the fatigue limit due to corrosion of aluminum bronze and monel metal in fresh water	137
1. Influence of stressless corrosion on the fatigue limit of aluminum bronze and monel metal	137
2. Influence of stress corrosion on the fatigue limit of aluminum bronze and monel metal	138
IV. Pits caused by corrosion of aluminum bronze with and without stress	139
1. Influence of the microstructure on the corrosion pitting of aluminum bronze	139
2. Pits caused by stressless corrosion of aluminum bronze in well water	140
3. Pits caused by stress corrosion at 1,450 cycles per minute	141
4. Pits caused by stress corrosion at 10,000 cycles per minute	142
5. Pits caused by stress corrosion at 500 cycles per minute	143
6. Pits caused by stress corrosion at 50 cycles per minute	143
7. Pits caused by stress corrosion at 10 cycles per minute	144
8. Pits caused by stress corrosion at 0.5 cycle per minute	144
9. Pits caused by stress corrosion at 5 cycles per hour	144
10. Pits caused by stress corrosion at 1 cycle per hour	145
11. Pits caused by stress corrosion at 1 cycle per day	145
V. Pits caused by corrosion of monel metal with and without stress	145
1. Pits caused by stressless corrosion of monel metal in well water	145
2. Pits caused by stress corrosion of monel metal at 1,450 cycles per minute	146
3. Pits caused by stress corrosion of monel metal at 5 cycles per hour	147
VI. Net and total damage due to stress corrosion of aluminum bronze and monel metal in well water	147
1. Types of diagrams representing net and total damage	147
2. Diagrams representing constant net damage	147
3. Diagrams representing constant total damage	149
VII. The process of stress corrosion of aluminum bronze and monel metal in aerated fresh water	150
VIII. Summary	151
IX. References	151

I. INTRODUCTION

In a recent paper [4] [1] the authors presented results of an investigation of the influence of cyclic stress on the corrosion pitting of steels in well water. The present paper considers the influence of both cyclic and steady stresses on the corrosion pitting of aluminum bronze and monel metal.

The lowering of the fatigue limit due to stress corrosion of these nonferrous alloys has been discussed in previous papers by the senior author [1, 2].[2] At that time, however, no investigation had been made of the forms and sizes of the corrosion pits and their correlation with the cyclic stress applied during corrosion. Later, through the courtesy of the Director of the U. S. Naval Engineering Experiment Station, Annapolis, Md., the specimens that had been used were made available for such a study. This paper gives the results of this investigation and correlates the forms and sizes of the corrosion pits with the resultant lowering of the fatigue limit.

II. MATERIALS AND METHOD OF INVESTIGATION

The chemical compositions of the cold-worked aluminum bronze and monel metal are given in table 1. Aluminum bronze of this composition consists of two or more microconstituents, one of them an alpha solid solution of aluminum and iron in copper. Monel metal consists almost entirely of a single microconstituent, a solid solution. Details of the annealing treatment of these alloys for relief of internal stress are given in table 2, and tensile properties are given in table 3.

TABLE 1.—*Chemical compositions*

[Averages of two or more determinations]

Alloy	Designation	Percentages of—								
		C	Ni	Cu	Al	Fe	Mn	P	S	Si
Aluminum bronze	L-6			87.7	10.5	2.64				
Monel metal	EP-8	0.16	67.5	29.5		1.76	0.95	0.015	0.017	0.05

TABLE 2.—*Annealing treatments for relief of internal stress*

Alloy	Designation	Temperature	Time	Cooled in—
		° *F*	*Minutes*	
Aluminum bronze	L-6	600	180	Furnace.
Monel metal	EP-8	800	180	Do.

TABLE 3.—*Tensile properties of alloys*

[Average of at least two determinations]

Alloy	Designation	Tensile strength	Johnson's limit	Proof stress (0.01%)	Elastic limit	Proportional limit	Elongation in 2 in.	Reduction of area
		lb/in.[2]	*lb/in.*[2]	*lb/in.*[2]	*lb/in.*[2]	*lb/in.*[2]	*Percent*	*Percent*
Aluminum bronze	L-6	101,900	41,900	42,800	36,900	32,900	16.6	17.7
Monel metal	EP-8	127,200	95,000	90,000	84,500	82,000	20.6	57.8

[1] Figures in brackets indicate the literature references at the end of this paper.

[2] The aluminum bronze was previously [2] designated KF-6. In this paper it is designated L-6 to avoid confusion with the effective stress concentration factor K_f.

Each specimen used in the investigation of the forms and sizes of corrosion pits had been subjected to corrosion under a given cyclic stress and cycle frequency for a given time, and had then been dried, oiled, and subjected to fatigue test in air. The lowering of the fatigue limit was used as a measure of the damage due to the corrosion. The water used in most of the corrosion experiments was the same well water that was used in the previous experiments with steels [4]. Some of the specimens, however, had been corroded in Severn River water, a brackish water with a saline content considerably higher than that of the well water.

In studying the corrosion pits, the specimens were examined visually on the outer surface, and typical specimens were photographed. The pits were also viewed in longitudinal sections cut approximately through the axes of the specimens. To insure a correct sectional view of the pits, the polishing was done in such a way as to minimize removal of corrosion products. For this purpose, the specimens were prepared by the method of dry polishing described in the previous paper [4]. Typical micrographs were than made. The magnification required was much higher ($\times 1000$) for these micrographs than for those shown in the previous paper [4].

III. LOWERING OF THE FATIGUE LIMIT DUE TO CORROSION OF ALUMINUM BRONZE AND MONEL METAL IN FRESH WATER

1. INFLUENCE OF STRESSLESS CORROSION ON THE FATIGUE LIMIT OF ALUMINUM BRONZE AND MONEL METAL

The influence of both stressless corrosion and stress corrosion on the fatigue limit is represented by figures 1 and 2. In these figures, abscissas represent durations of corrosion and ordinates represent the fatigue limits as determined by subsequent tests without corrosion. From these diagrams are derived diagrams of several types representing the relation between stress, corrosion time, and cycle frequency for constant total and net damage. The derived diagrams will be discussed (section VI) after consideration of the results of examination of the pitted specimens.

During the stressless corrosion of steel, as shown in the previous paper [4], the fatigue limit decreases rapidly at first, but the rate of decrease gradually slackens and eventually becomes very slow. Curves representing such a variation of the fatigue limit have been termed "retarded-damage" curves. Aluminum bronze and monel metal, however, give stressless-corrosion curves of a very different type. As shown in figures 1 and 2, each of these curves is nearly horizontal at first and remains nearly horizontal throughout a long corrosion period. The slope, however, gradually increases and eventually becomes steep. Curves of this form, termed "accelerated-damage" curves, have also been obtained with stainless steel and nickel [2].

As shown in the preceding paper [4], the rate of corrosion of steel in well water is cathodically controlled (except possibly at first); that is, the rate is determined by the conditions at the cathodes. Corrosion of aluminum bronze, monel metal, stainless steel, and nickel, in well water, however, is anodically controlled; that is, the

rate of corrosion is determined by the conditions at the anodes. The evidence indicates, therefore, that cathodically controlled corrosion tends to give stressless-corrosion curves of the retarded-damage type, and that anodically controlled corrosion tends to give stressless-corrosion curves of the accelerated-damage type, such as those in figures 1 and 2. Space is not available for discussion of the types of stressless-corrosion curves obtained when the corrosion rate depends on conditions at both cathodes and anodes.

The results of corrosion generally are more erratic when the control is anodic than when it is cathodic. With anodic control, specimens corroded under apparently identical conditions may differ greatly in the resultant damage as measured by the lowering of the fatigue limit. The scatter of experimental points representing stressless corrosion, consequently, is much greater in figures 1 and 2 than in graphs representing stressless corrosion of steel in well water [4]. In an accelerating corrosion process, as in many other accelerating processes, the results are less reproducible than in a process causing retarded damage.

Stressless corrosion of aluminum bronze for more than 2 years (fig. 1) reduced the fatigue limit to less than two-thirds the original value. If the acceleration of damage should persist with increase in the corrosion time, the fatigue limit evidently would be reduced to zero in about 4 years, thus implying that corrosion pits had reached the axis of the specimen. Before this, however, the rate of lowering of the fatigue limit might begin to slacken, thus causing a reversal of the stressless-corrosion curve. (Such a delayed reversal would not affect the derived diagrams representing constant net and total damage).

Stressless corrosion of monel metal (fig. 2) would be better represented by a widening band than by a single curve. The curve as drawn takes an approximate mean course. In drawing the curve, consideration has been given not only to the experimental points representing stressless corrosion but also to the three points representing corrosion under stresses of 8,000, 10,000, and 12,000 lb/in.^{2} at 1,450 cycles per minute.

2. INFLUENCE OF STRESS CORROSION ON THE FATIGUE LIMIT OF ALUMINUM BRONZE AND MONEL METAL

The stress-corrosion curves for aluminum bronze and monel metal, unlike many of the curves for steels [4], descend at an increasing rate. The scatter of experimental points generally is greater for a stress-corrosion curve than for a stressless-corrosion curve. The scatter is much greater for monel metal than for aluminum bronze. Even though several experiments were made with monel metal at a number of the stresses, the scatter of results is so great that the curves could not be accurately established. Some experiments caused fatigue failure during corrosion, whereas other experiments at the same stress and for about the same corrosion time caused little damage. Such results indicate that an ideal stress-corrosion curve would show an accelerated increase of slope and final abrupt drop to zero ordinate. The curves are so drawn in figures 1 and 2. Each curve is drawn so as to follow an approximate mean course.

Under simultaneous corrosion and cyclic stress, the lowering of the fatigue limit evidently is more rapid than under stressless corrosion.

For a given cycle frequency, the rapidity of descent of a curve increases with the cyclic stress; for a given stress, the rapidity of descent increases with the cycle frequency. Even at very low cycle frequencies (ranging from 5 cycles per hour to 1 cycle per day), however, sufficiently high stress and sufficiently long duration of corrosion may cause much net damage (figs. 1 and 2). The quantitative relation between stress, cycle frequency, corrosion time, and either net or total damage [3] is discussed in section VI.

IV. PITS CAUSED BY CORROSION OF ALUMINUM BRONZE WITH AND WITHOUT STRESS

1. INFLUENCE OF THE MICROSTRUCTURE ON THE CORROSION PITTING OF ALUMINUM BRONZE

The position and form of corrosion pits in steels, in aerated well water, are little affected by the microstructure [4]. The pits are free to assume the forms determined by the corrosion conditions. Pitting in the aluminum bronze used in these experiments, however, was much influenced by the microstructure. This is illustrated by the positions and forms of the pits caused by corrosion of a specimen in salt water. The water used for this purpose was Severn River water, which generally has about one-sixth the saline content of sea water. The specimen was corroded under stress of 60,000 lb/in.^{2} at 1 cycle per day for 16 days. Surface views are shown in figure 9 and longitudinal sections are shown in figure 10.

Views *A* and *B* of figure 9 show the surface of this corroded specimen after slight polishing with magnesium oxide. The polishing was merely enough to remove most of the corrosion products from the cathodic areas and reveal the pits. The dark areas, shown in view *A* at 100 magnification and in view *B* at 500 magnification, are pits filled with corrosion products. These corroded areas are in the delta phase of the alloy. The nearly white, uncorroded areas are in the alpha solid solution. The alpha phase evidently is cathodic to the delta phase and to the associated iron-rich constituent.

After considerable additional polishing with magnesium oxide, the surface appears as shown in views *C* and *D*. The cathodic areas are now practically free from corrosion products, and the dendritic structure of the alloy is clearly revealed. The corrosion pits in this alloy generally are not free to assume the roughly hemispherical or saucer-like forms that are found after corrosion of steels [4]. The pits in the aluminum bronze are bounded by the frondlike crystallites of the cathodic alpha phase.

This influence of the microstructure is revealed not only by the surface views but also by the sectional views (fig. 10). Whereas pits in steel tend to deepen perpendicularly to the surface, the pits in the aluminum bronze follow the interstices between adjacent crystallites of the alpha solid solution. Many of the pits thus take an oblique direction (fig. 10). This influence of the microstructure is found not only after corrosion in salt water but also after the much slower corrosion in fresh water. Numerous instances of this influence will be found in the micrographs now to be considered.

[3] Total damage due to stress-corrosion is measured by the vertical distance of a stress-corrosion curve (at any point) below the origin of the curve. Net damage is the difference between the total damage and the damage that would be caused in the same time by stressless corrosion.

Although the entire delta phase tends to be anodic during corrosion in either salt water or well water, pitting in well water is not continuous throughout the delta constituent. Some parts of this phase evidently are more anodic than others. Very little information about the forms and distribution of the corrosion pits, therefore, can be obtained by examining the darkened anodic areas on the surface. In this investigation, consequently, attention has been confined almost entirely to longitudinal sections. Results of the investigation are shown in figures 11 to 26, inclusive.

2. PITS CAUSED BY STRESSLESS CORROSION OF ALUMINUM BRONZE IN WELL WATER

Although results of stressless corrosion of five specimens of aluminum bronze in well water are shown in figure 1, only four of these specimens were available for examination of the corrosion pits. Sectional views of these specimens are shown in figures 11 and 12.

Corrosion for 66 days (view *A*, fig. 11) has caused numerous small pits. The forms of most of these pits evidently have been influenced greatly by the duplex microstructure of the alloy. A few of the pits are roughly hemispherical, like pits caused by stressless corrosion of steel. After corrosion for 88 days, the pits shown in view *B* of figure 11 are smaller than those shown in view *A*. The lowering of the fatigue limit, however, was slightly greater after corrosion for 88 days than after corrosion for 66 days (fig. 1). The sections shown in figure 11, *B*, therefore, probably do not traverse the regions of deepest pitting on this specimen. As the variation of pitting in a specimen generally is greater when the process is anodically controlled than when it is cathodically controlled, sectional views of the most deeply pitted regions can be obtained less easily with aluminum bronze (and monel metal) than with steels.

After corrosion for much longer than 66 days (fig. 12), the pits have increased greatly in size, and the influence of microstructure on the forms of the pits evidently has become less prominent. Although most of the pits here shown have the saucerlike form found after long corrosion of steels, some of the pits show the confining influence of the alpha phase. The lowering of the fatigue limit (fig. 1) was much greater after corrosion for 640 and 1,021 days than after 66 and 88 days. This must be attributed largely to the influence of corrosion time on the size of the pits. The slightness of the effect of the pits shown in figure 11 evidently is due to their small size [4].

After corrosion of steel in well water for only 4 and 4.7 days (figs. 9 and 33 of the previous paper [4]),[4] the pits were much larger and the lowering of the fatigue limit was much greater than after corrosion of aluminum bronze in well water for 66 and 88 days (fig. 11).

Values of the effective stress concentration factor, K_F, obtained by stressless corrosion of steels,aluminum bronze, and monel metal are assembled in table 4. The values for the breadth and depth of pits have been taken from the sectional views of the pits that probably would cause the highest values of K_F. These were obtained from figures 9, 10, and 33 of the first paper and figures 11, 12, 22, 23, and 24 of the present paper. The values of K_F were obtained from the corresponding diagrams, one of which is not found in either of these two papers.[5]

[4] In such comparison, allowance must be made for the fact that the magnification of the sectional views in this paper is 4 to 20 times as great as the magnifications used in the previous paper.

[5] The diagram for steel AX-W-10 is in a previous paper [3] by the senior author.

TABLE 4.—*Sizes of stressless-corrosion pits and corresponding values of the effective stress-concentration factor, K_F*

Metal	Designation	Water	Days	Size of pit		Ratio 5:6	K_F
				Breadth	Depth		
1	2	3	4	5	6	7	8
				Inches	*Inches*		
Steel	AX-W-10	Well	2	0.0024	0.0014	1.71	1.09
Do	AX-W-10	do	4.7	.006	.0024	2.5	1.35
Do	IW-W-10	do	10	.010	.003	3.3	1.47
Do	AX-W-10	do	47	.012	.007	1.71	1.61
Do	IW-W-10	do	50	.013	.005	2.6	1.71
Do	IW-W-10	do	100	.025	.011	2.27	1.83
Aluminum bronze	L-6	do	66	.004	.002	2.0	1.05
Do	L-6	do	640	.00055	.0004	1.37	1.54
Do	L-6	do	1,021	.00065	.0004	1.62	1.76
Monel metal	EP-8	do	122	.0008	.0004	2.0	1.10
Do	EP-8	do	157	.0009	.00045	2.1	1.10
Do	EP-8	do	170	.0011	.0009	1.22	1.20
Do	EP-8	do	170	.0005	.0011	0.45	1.20

For the steels, the value of K_F evidently tends to increase with increase in the breadth of the corrosion pits. A similar relationship is found for aluminum bronze and monel metal in well water. If table 4 were arranged solely in the order of increasing breadth of the corrosion pits, however, this would not be the order of increasing value of K_F. The values obtained with aluminum bronze in well water for 640 and 1,021 days are too high to be in the order of increasing breadth of corrosion pits. (Some of the values for monel metal also are too high for such an arrangement.)

By plotting values of the breadth of a pit as abscissas and corresponding values of K_F-1 as ordinates, the points obtained with steels in well water would give a curve starting from the origin of coordinates and rising at a decreasing rate. (Such a curve is also obtained by fatigue tests with specimens having notches of the same form but differing in size.) The points obtained with aluminum bronze in well water for 640 and 1,021 days would be above this curve. Probable reasons for this are found by considering the ratios of breadth to depth of the pits. For the steels in well water, the mean value of this ratio is about 2.3; that is, the pits generally are roughly hemispherical. For the aluminum bronze, after corrosion for 640 and 1,021 days, the ratio is only about 1.5. The relatively high values of K_F obtained with these specimens, therefore, may be attributed in part to the fact that the relative depth of the pits is greater in this alloy than in steels. They may be attributed in part, however, to the angularity of some of the corrosion pits, the angularity resulting from the duplex microstructure of the aluminum bronze. The differences in the values of K_F in table 4 evidently are due not only to differences in size but also to differences in the forms of the pits.

3. PITS CAUSED BY STRESS CORROSION AT 1,450 CYCLES PER MINUTE

In studying the influence of cyclic stress on the forms and sizes of corrosion pits, the pits caused by stressless corrosion will be compared with the pits caused by simultaneous corrosion and cyclic stress. Attention will be given first to specimens that have been subjected

to various stresses with a frequency of 1,450 cycles per minute. Typical sectional views are shown in figure 13. These were obtained after corrosion for short times.

The forms of nearly all these pits have been greatly influenced by the duplex microstructure. The pits, therefore, are similar in form to many of the pits caused by stressless corrosion (figs. 11 and 12). The cyclic stress, however, evidently has had much influence on the size of the pits. After corrosion at 30,000 lb/in.^{2} for only 1.8 days, the pits (view *A*) are nearly as large as those found after stressless corrosion for 66 days (fig. 11, *A*). After corrosion at 25,000 lb/in.^{2} for 6 days (fig. 13, *B*), the pits are nearly as large as those found after stressless corrosion for 640 days (fig. 12, *A*). After corrosion at 20,000 lb/in.^{2} for 11 days (fig. 13, *C*), some of the pits are larger than those found after stressless corrosion for 1,021 days (fig. 12, *B*). Moreover, the ratio of depth to breadth, for some of the pits shown in figure 13, is greater than for the pits found after stressless corrosion. Although no distinct fissures or crevices (such as those found after corrosion of steels at this cycle frequency) are found in figure 13, the cyclic stress probably has increased the sharpness of the salients in this alloy. Cyclic stress evidently influences the size and form of the corrosion pits in aluminum bronze as well as in steels.

To facilitate comparison between the corrosion pits and the corresponding lowering of the fatigue limit, three different symbols are used for the experimental points in figures 1 and 2. A triangle is used when the sectional view gives distinct evidence of fissures; a square is used when no fissures are visible; a circle is used when the pits have not been examined.

The pits in any specimen tend to vary more widely when the corrosion process is anodically controlled than when it is cathodically controlled. Correlation between a sectional view of corrosion pits and the lowering of the fatigue limit, therefore, is less satisfactory for aluminum bronze than for steels. If the pits shown in figure 13 were the most advanced pits in these three specimens, these views would be in the order of increasing damage. Actually, however, they are in the order of decreasing damage (fig. 1). The pit at which fatigue fracture started probably was more advanced than any of the pits found in a sectional view. The resultant damage evidently depended on the increased size and sharpness of a few of the most advanced pits, not visible in figure 13.

The apparent absence of damage (fig. 1) due to the pits shown in figure 13 (*C*) may be attributed to a locally high fatigue limit for this specimen. The pits shown in figure 13 (*C*), although they appear large at this high magnification, are smaller than pits caused by stressless corrosion of steel for a much shorter time in well water (fig. 9 of the preceding paper [4]).

4. PITS CAUSED BY STRESS CORROSION AT 10,000 CYCLES PER MINUTE

Corrosion with a cycle frequency of 10,000 per minute at 30,000 lb/in.^{2} for 0.3 day has caused the pits shown in view *A* of figure 14. One of the pits here shown is about as large as any of those found after stressless corrosion for 66 days (fig. 11, *A*). The pits shown in figure 14 (*A*), moreover, generally have greater relative depth than those found after stressless corrosion. Some of the pits, especially

those extending in oblique direction, are very sharp. From one of these a crack has started. A crack such as this evidently caused the fatigue failure during the corrosion stage (fig. 1).

5. PITS CAUSED BY STRESS CORROSION AT 500 CYCLES PER MINUTE

Corrosion with a cycle frequency of 500 per minute has caused the pits shown in views *B*, *C*, and *D* of figure 14. After corrosion at 60,000 lb/in.^{2} for 0.06 day (view *B*), some pits are nearly as large as those found after corrosion for 640 days without stress (fig. 12, *A*). The cyclic stress, moreover, evidently has caused the sharp projections generally extending in an oblique direction. To sharp salients such as these may be attributed the rapid decrease of the fatigue limit during corrosion (fig. 1). As this specimen did not fail by fatigue during corrosion, the cracks starting from some of the sharp salients (fig. 14, *B*) may have started during the subsequent fatigue test.

The sharp projections extending inward from the corrosion pits shown in views *A* and *B* of figure 14 probably are crevices and fissures[6] similar to those caused by stress corrosion of steels. In aluminum bronze, however, the crevices and fissures generally are not free to advance in a plane perpendicular to the axis of the specimen, as they do in steels. The planes of extension, both inward and along the surface of the specimen, are determined largely by the direction of the crystallites of the alpha phase. All these fissures probably have started at blunt pits and have extended both inward and along the surface. A sectional view, in aluminum bronze as in steel [4], however, may not traverse the origin of a fissure.

Corrosion at 40,000 lb/in.^{2} for 0.7 day has caused the pits shown in figure 14 (*C*). As the corrosion caused considerable damage (fig. 1), this damage must be attributed to larger or sharper pits than those revealed in figure 13 (*C*). A similar discrepancy exists between the damage and the pits shown in figure 13 (*D*).

6. PITS CAUSED BY STRESS CORROSION AT 50 CYCLES PER MINUTE

Corrosion with a cycle frequency of 50 per minute has caused the pits shown in figure 15. After corrosion at 50,000 lb/in.^{2} for 1.1 days (view *A*), some pits are nearly as large as those found after stressless corrosion for 640 days (fig. 12, *A*). The cyclic stress, moreover, has caused sharp oblique projections, as shown in some of the sections. To such projections, which probably are sections of fissures, may be attributed the rapid lowering of the fatigue limit (fig. 1). These fissures probably started at blunt pits, such as those shown in figure 15 (*A*), and extended both along the surface and inward.

After corrosion at 35,000 lb/in.^{2} for 6.7 days (fig. 15, *B*), some of the pits are larger than those found after stressless corrosion for 1,021 days (fig. 12, *B*). One of the pits shown in figure 15 (*B*) has extended inward in an oblique direction and has incipient sharp projections. To such projections and to the increased size of the pits may be attributed the lowering of the fatigue limit, as shown in figure 1.

After corrosion at 20,000 lb/in.^{2} for 52 days, the pits are larger than those found after stressless corrosion for 66 days (fig. 11, *A*). Because

[6] Crevices, as defined in the previous paper [4], are merely fissures that have not advanced far enough to be visible on the surface of the specimen at low magnification. No attempt will be made in this paper to distinguish between crevices and fissures.

of the small size of these pits (only about 0.0005-inch diameter), their effect on the fatigue limit was inappreciable (fig. 1).

7. PITS CAUSED BY STRESS CORROSION AT 10 CYCLES PER MINUTE

Corrosion with a cycle frequency of 10 per minute has caused the pits shown in figure 16. After corrosion at 40,000 lb/in.^{2} for 13 days (view *A*), some pits are as large as those found after stressless corrosion for 1,021 days (fig. 12, *B*). Sharp projections extend obliquely inward. The largest of these pits is the origin of sharp projections, which probably are sections of fissures. This view, unlike those in figures 14 (*A*) and (*B*) and 15 (*A*), shows the blunt pit from which the fissures started. The fissures have caused considerable damage (fig. 1).

After corrosion at 30,000 lb/in.^{2} for 26 days (fig. 15, *B*), the pits are about as large as those found after stressless corrosion for 66 days (fig. 11, *A*). Incipient sharp projections, such as those here shown, may account for the considerable decrease in the fatigue limit (fig. 1).

8. PITS CAUSED BY STRESS CORROSION AT 0.5 CYCLE PER MINUTE

Corrosion with a cycle frequency of 0.5 per minute has caused the pits shown in figure 17. After corrosion at 60,000 lb/in.^{2} for 6 days, the pits found in view *A* are very small; one of the blunt pits, however, is relatively deep. A few sharp pits extending inward from the surface probably are sections of fissures which started at blunt pits not shown in this view. This effect of the cyclic stress accounts for the considerable damage shown in figure 1.

After corrosion at 50,000 lb/in.^{2} for 13 days (fig. 18, *B*), some of the pits are as large as those found after stressless corrosion for 640 days (fig. 12, *B*). Sharp pits, moreover, extend obliquely inward from the surface. The blunt origins of these sharp pits are not traversed by this section. The lowering of the fatigue limit of this specimen was not as great as would be expected.

9. PITS CAUSED BY STRESS CORROSION AT 5 CYCLES PER HOUR

Corrosion with a cycle frequency of 5 per hour has caused the pits shown in figures 18 and 19. After corrosion at 60,000 lb/in.^{2} for 15 days (fig. 18, *A*), the pits are as large as those found after stressless corrosion for 1,021 days (fig. 12, *B*). One of the sections, moreover, shows an oblique pit, which probably is an incipient fissure. The increased size and sharpness of the pits, due to the cyclic stress, account for the lowering of the fatigue limit (fig. 1).

After corrosion at 50,000 lb./in.^{2} for 24 days (fig. 18, *B*), the pits are considerably larger than those found after stressless corrosion for 1,021 days. Nevertheless, the pits are much smaller than those caused by corrosion of steel in well water for 10 days (fig. 9, *D* of the first paper [4]). The sharp oblique projections from one of the pits in figure 18 (*B*) probably are incipient fissures. The lowering of the fatigue limit was about the same for this specimen as for the specimen represented in figure 18 (*A*).

After corrosion at 30,000 lb/in.^{2} for 339 days, the pits shown in figure 19 are only about as large as those found after stressless corrosion for 66 days (fig. 11, *A*). This specimen, therefore, probably contained much larger pits than those traversed by the section.

Numerous sharp projections, however, extend inward from pits and from the surface of the specimen. To these sharp fissures may be attributed the great decrease of the fatigue limit (fig. 1).

10. PITS CAUSED BY STRESS CORROSION AT 1 CYCLE PER HOUR

Corrosion with a cycle frequency of 1 per hour has caused the pits shown in figure 20. After corrosion at 60,000 lb/in.^{2} for 20 days (view *A*), the pits are larger than those found after stressless corrosion for 66 days (fig. 11, *A*) but no fissures are visible. The effect on the fatigue limit, as shown in figure 1, was small. About the same effects on the pits and on the fatigue limit were caused by corrosion at 40,000 lb/in.^{2} for 73 days (fig. 20, *B*). Corrosion at the same stress for a much longer time (fig. 20, *C*), however, caused sharp projections, which probably are sections of fissures. The development of these fissures has caused rapid descent of the curve of decrease of the fatigue limit (fig. 1).

After corrosion at 35,000 lb/in.^{2} for 374 days (fig. 20, *D*), no large pits are found. This specimen, however, probably contained much larger pits than those traversed by the sectional view. Sharp fissures extend obliquely inward from the surface and from some of the small pits. To the sharp fissures must be attributed the great decrease in the fatigue limit (fig. 1). (In a few more days, this specimen probably would have failed during corrosion.)

11. PITS CAUSED BY STRESS CORROSION AT 1 CYCLE PER DAY

Corrosion with 1 cycle per day has caused the pits shown in figure 21. After corrosion at 60,000 lb/in.^{2} for 28 days (view *A*), the pits are nearly as large as those found after stressless corrosion for 640 days (fig. 12, *A*). The pits have lowered the fatigue limit slightly. After corrosion at 40,000 lb/in.^{2} for 151 days, sharp projections are found (view *B*, fig. 21). In some regions, these sharp salients are intermingled with portions of the alpha phase. Pits of the same type, moreover, are found after corrosion at the same stress for 240 days (view *C*). The sharp salients, together with interspersed portions of the alpha phase, form pits larger than those found after stressless corrosion for 1,021 days (fig. 12, *B*). The great effect of such pits on the fatigue limit is illustrated by the descent of the curve in figure 1.

V. PITS CAUSED BY CORROSION OF MONEL METAL WITH AND WITHOUT STRESS

1. PITS CAUSED BY STRESSLESS CORROSION OF MONEL METAL IN WELL WATER

The distribution of corrosion pits on the specimens of monel metal was far from uniform. The pits generally were found in clusters. Through such a cluster in each specimen a section was cut and polished for microscopic examination. After the pits in this section were photographed, the section generally was polished again so as to remove about 0.001 inch of metal. The removal of this thin layer of metal generally revealed other corrosion pits. In this manner, photographs of several sections sometimes were obtained. In assembling the photographs of each specimen, no attempt was made to distinguish between the different sections.

Only a few of the many specimens of monel metal that were subjected to corrosion were available for examination of the corrosion pits. The pits were examined in only three of the nine specimens (fig. 2) that had been subjected to stressless corrosion. Sectional views of these three specimens are shown in figures 22, 23, and 24.

After stressless corrosion for 122 days (fig. 22), the pits are small. Although some pits appear large at this high magnification, they are much smaller than the pits caused by corrosion of steel for only a few days (fig. 9 of the previous paper [4]). As monel metal is essentially a single-phase alloy, the pits are little influenced by the microstructure. The pits shown in figure 22, consequently, are roughly hemispherical or saucerlike, resembling in this respect the pits caused by brief corrosion of steel. Because of their bluntness and small size, these pits have had little effect on the fatigue limit (fig. 2).

After stressless corrosion for 157 days (fig. 23), some of the pits are broader, but no deeper, than after corrosion for 122 days. Some of the pits resemble in form and size the pits shown in figure 22. The fatigue limit of this specimen, moreover, was about the same as that of the specimen corroded for 122 days.

Larger pits were found after stressless corrosion for 170 days (fig. 24). Sharp projections, moreover, have formed on some of the pits; other sharp pits have extended inward from the surface. As these sharp pits were formed without stress, they probably are rootlike projections rather than transverse fissures. Rootlike pits may be formed in either steels or nonferrous metals under the influence of concentration gradients of oxygen or of an anion or cation. The fatigue limit of this specimen was lower than that of all the other stressless-corrosion specimens of this alloy, even specimens that had been corroded for much longer times. Typical stressless-corrosion pits in monel metal, therefore, probably resemble in form those seen in figures 22 and 23 rather than those seen in figure 24. This conclusion is supported by results of examination of pits caused by stress corrosion.

2. PITS CAUSED BY STRESS CORROSION OF MONEL METAL AT 1,450 CYCLES PER MINUTE

Corrosion with a cycle frequency of 1,450 per minute has caused the pits shown in figure 25. After corrosion at 30,000 lb/in.^{2} for 10 days, only a few very small pits are visible even at this high magnification (view *A*). They probably are not much larger than if the specimen had been corroded without stress. Rounded pits of this size (about 0.0001-inch diameter) have practically no effect on the fatigue limit [4]. The fatigue limit, however, was not much more than half that of an uncorroded specimen (fig. 2). Some pits in this specimen, therefore, probably were much larger or sharper than if the specimen had been corroded without stress. This is one of the few experiments (at high frequency) that were stopped in the midst of the nearly vertical descent of the stress-corrosion curve. As this steep descent probably was due to the rapid development of a single fatigue crack from one of the corrosion pits, the chance of finding the rounded pit from which the crack started was very small.

After corrosion at 10,000 lb/in.^{2} for 331 days (fig. 25, *B*) some of the pits are much larger than those caused by stressless corrosion for 170 days (fig. 24). Most of this difference in size probably is due to the cyclic stress. The great decrease of the fatigue limit caused by

stress corrosion of this specimen (fig. 2), however, probably should be attributed to the development of a crevice, fissure, or fatigue crack from a sharp projection, such as the one pointing inward from the largest pit in figure 25. As shown in the previous paper [4], the effective stress-concentration factor increases with both the size and the sharpness of the pits.

3. PITS CAUSED BY STRESS CORROSION OF MONEL METAL AT 5 CYCLES PER HOUR

Corrosion with a cycle frequency of 5 per hour has caused the pits shown in figure 26. After corrosion at 45,000 lb/in.^{2} for 123 days (view *A*), the pits are larger than those found after stressless corrosion for about the same time (fig. 22), and are about as large as those found after stressless corrosion for 170 days (fig. 24). Some of the pits in figure 26 (*A*), moreover, have small sharp projections, which probably represent incipient fissures. The effect of the cyclic stress on the size and form of the pits, however, was not great enough to cause much net damage (fig. 2). Corrosion for a little longer time probably would have caused rapid development of transverse fissures and rapid descent of the corresponding curve in figure 2.

Corrosion at 40,000 lb/in.^{2} for 148 days has caused the pits shown in view *B* of figure 26. Although the corrosion time was longer for this specimen than for the specimen represented in view *A*, the moderate increase in time would not be expected to compensate for the decrease in stress. As would be expected, therefore, the pits shown in view *B* appear less damaging than those shown in view *A*. Although some of the pits shown in view *B* are broader than those in view *A*, they are shallower and show no evidence of incipient fissures. The pits evidently have caused very little net damage (fig. 2).

VI. NET AND TOTAL DAMAGE DUE TO STRESS CORROSION OF ALUMINUM BRONZE AND MONEL METAL IN WELL WATER

1. TYPES OF DIAGRAMS REPRESENTING NET AND TOTAL DAMAGE

From figures 1 and 2, other diagrams have been derived to represent the interrelation between stress, corrosion time, cycle frequency, and total number of cycles, for constant net and total damage. These diagrams are shown in figures 3 to 8. Diagrams of the types shown in figures 3, 4, and 5 are shown in previous papers by the senior author [1, 2]. Since that time, however, stress-corrosion experiments with aluminum bronze were continued so as to remove all uncertainty as to the forms of these diagrams.[7] The additional information thus made available, together with the information obtained by examination of the corrosion pits, leads to definite conclusions as to the process of stress corrosion of these alloys.

2. DIAGRAMS REPRESENTING CONSTANT NET DAMAGE

The diagrams in figures 3, 4, and 5 represent the relation between stress, corrosion time, cycle frequency, and total number of cycles, for 15 percent net damage. This amount of net damage is represented

[7] These experiments were made with the cooperation of W. C. Stewart, Metallurgist of the U. S. Naval Engineering Experiment Station, with the permission of the Director of the Station.

by the broken line below each stressless corrosion line in figure 1. Each of these broken lines is drawn so that the ordinate at any point is 85 percent of the corresponding ordinate of the stressless-corrosion line. The intersections of the broken lines with the stress-corrosion lines are used in deriving the diagrams in figures 3, 4, and 5.

In figures 3 and 4, the diagrams may be considered to represent plan views of three-dimensional diagrams, in which the "vertical" coordinates (perpendicular to the plane of figs. 3 and 4) represent stresses in the corrosion stage. The curves in figures 3 and 4 thus represent contour lines on the surface of the three-dimensional diagrams. Complementary views of the three-dimensional diagrams for these two alloys are shown in figure 5. Each diagram in this figure represents lines on the surface as viewed in the direction of the arrow in figures 3 and 4. Diagrams *A* and *C* of figure 5 represent views of constant-frequency lines, which are seen in figures 3 and 4 as straight diagonal lines. Diagrams *B* and *D* of figure 5 show lines representing constant numbers of cycles, which are seen as constant-ordinate lines in figures 3 and 4. Diagrams of this type have not been considered in previous papers.

The form of the diagram for aluminum bronze (fig. 3) is well established by means of the experimental points. The curves as drawn, however, are not based only on the positions of these experimental points. The curves in this diagram and in the complementary diagrams (*A* and *B* of fig. 5) are so drawn that the three diagrams represent correlated views of a three-dimensional diagram. The diagram in figure 3 has thus been extended to lower stresses and cycle frequencies than those used in actual experiments. The extrapolated portion of the diagram may be considered qualitatively correct.

The form of the corresponding diagram for monel metal (fig. 4) also is well established by means of the experimental points. The diagram for monel metal evidently is qualitatively similar to the diagram for aluminum bronze. Drawn in accordance with the evidence and in proper correlation with the complementary diagrams (*C* and *D*, fig. 5), the diagram in figure 4 may be considered qualitatively correct.

The diagrams in figures 3 and 4 are very different from the corresponding diagram obtained with steels in well water, figure 6, of the previous paper [4]. Whereas all the curves in that diagram, as they extend to the right, approach a horizontal direction, all the curves in figures 3 and 4 approach a vertical direction. This difference in form is due chiefly to the difference in form of the curves representing stressless corrosion. When the stressless-corrosion curve is of the retarded-damage type, the derived diagram tends to be of the form obtained with steels in well water. When the stressless-corrosion curve is of the accelerated-damage type (figs. 1 and 2), the derived diagram tends to be of the form shown in figures 3 and 4.[8] The great difference in form of these diagrams representing constant net damage, therefore, may be attributed to the difference in the type of control of the rate of corrosion. Anodic control tends to give a diagram of the form shown in figures 3 and 4.

All the diagrams in figure 5 are similar in form. The lines representing low frequencies (in diagrams *A* and *C*) and the lines representing small numbers of cycles (in diagrams *B* and *D*) are nearly straight,

[8] The diagrams representing constant net damage thus resemble qualitatively the diagrams representing constant total damage, figures 6 and 7.

as are all the lines in the corresponding diagrams for steels in well water [4]. With increase in the cycle frequency or in the total number of cycles, however, the curvature of the lines increases and the slope decreases.

From the slope of a low-frequency line, an estimate can be made of the influence of stress on the rate of net damage. As shown in the preceding paper [4], the linear logarithmic relationship may be represented by the equation $R=CS^n$, in which R represents the mean rate of net damage, S the stress, and C and n are constants. The exponent n is the cotangent of the angle of slope of the line. As indicated by the slopes of the lines representing 1 cycle per day in figure 5, the value of n (for this cycle frequency) is 3.5 for aluminum bronze and about 2.1 for monel metal. The rate of net damage evidently varies as a power of the stress. As represented by the curved lines in figure 5, however, the power exponent increases with the stress and with the cycle frequency. In this respect, the diagrams for aluminum bronze and monel metal differ from the corresponding diagram for steels (fig. 5 of the preceding paper [4]).

3. DIAGRAMS REPRESENTING CONSTANT TOTAL DAMAGE

A diagram to represent constant total damage may be derived from families of curves of the type shown in figures 1 and 2 by utilizing the coordinates of intersections of such curves with a given horizontal line. An index of the total damage corresponding to this horizontal line is the effective stress concentration factor, K_F, obtained by dividing the fatigue limit of an uncorroded specimen by the fatigue limit of the corroded specimen [4].

Two-dimensional diagrams representing constant total damage may be of several types, representing different views of a three-dimensional diagram. In figures 6, 7, and 8 are diagrams of three types, corresponding to the three types of diagrams representing constant net damage (figs. 3, 4, and 5). The diagrams in figures 6 and 7 are plan views of three-dimensional diagrams for aluminum bronze and monel metal, respectively. In figure 8, the diagrams represent views in the direction of the arrows in figures 6 and 7. In diagrams A and C of figure 8, the lines are constant-frequency lines; in diagrams B and D, each line represents the indicated total number of cycles. The total damage on which each diagram is based is indicated by the value of K_F. At the right of each of the diagrams in figures 6 and 7 and at the left of each diagram in figure 8, is a straight boundary line which represents the time necessary to reach the indicated value of K_F through stressless corrosion. The course of each curve represents the influence of cycle frequency (or number of cycles) in shortening the time necessary to cause constant total damage.

In the plan diagram for aluminum bronze (fig. 6), each curve approaches a vertical direction. The curves, however, do not approach the same vertical asymptote. With the possible exception of curves representing stresses less than about 25,000 lb/in.^{2}, none of the curves approaches the time boundary of the diagram, the line representing stressless corrosion. A diagram of this form evidently indicates that the influence of stress on corrosion would be about the same for all cycle frequencies less than about 1 cycle per day. The evidence in figure 6 thus implies that the corrosion pitting of aluminum bronze is accelerated by steady stress. The relation between steady stress

and time for constant net damage evidently is represented approximately by the curve representing 1 cycle in figure 8 (*B*).

The plan diagram for monel metal (fig. 7) is qualitatively similar to the diagram for aluminum bronze. The diagram for monel metal, however, has not been extended to sufficiently low cycle frequencies to determine whether or not all the curves approach the time boundary of the diagram. If they do, steady stress (at least up to 80,000 lb/in.^{2}) has no appreciable effect on the corrosion pitting of this alloy. Comparison of figures 6 and 7, however, shows conclusively that the influence of steady stress is much less for monel metal than for aluminum bronze. The same conclusion would be reached by comparing the diagrams for these two alloys in figure 8.

Diagrams representing constant total damage for steels in fresh water (not shown) differ greatly from the diagram for aluminum bronze (fig. 6). In a plan diagram for steels, all the curves approach the time boundary, thus indicating that steady stress (at least up to the maximum value used in the investigation) has no appreciable effect on the corrosion pitting of steel in fresh water. As shown in the preceding paper [4], an influence of steady stress on corrosion would not be expected when the rate of corrosion is cathodically controlled, but is possible when the rate is anodically controlled.

VII. THE PROCESS OF STRESS CORROSION OF ALUMINUM BRONZE AND MONEL METAL IN AERATED FRESH WATER

The evidence presented in previous sections indicates that cyclic stress tends to increase the size and sharpness of corrosion pits in aluminum bronze and monel metal. The effect tends to increase with the stress and with the cycle frequency. The effect of stress on the corrosion pitting of these two alloys thus resembles qualitatively the effect of stress on the corrosion pitting of steels. For steels in well water [4], however, the accelerating influence of cyclic stress is caused chiefly by an increase in the permeability of the coating of corrosion products. Cyclic stress increases the permeability in both cathodic and anodic regions and tends to shift the process from a cathodically controlled toward an anodically controlled process. An increase in the stress or the cycle frequency thus opposes the tendency of the process to shift, with time, toward purely cathodic control. As the corrosion of aluminum bronze and monel metal is an anodically controlled process, the rate of corrosion pitting depends not only on the initial stress but also on the enhanced stress due to stress concentration around the corrosion pits. The rate of pitting thus tends to be accelerated. As with any accelerating process, the results are less reproducible than the results obtained by stress corrosion of steels in well water [4].

The accelerating influence of steady stress on the corrosion pitting of aluminum bronze may be due to an increase in the solution pressure, especially in the regions of high stress concentration around corrosion pits. This influence of stress is superposed on the influence of an increased permeability of the anodic coating owing to the cyclic stress.

The difference in behavior of monel metal and aluminum bronze under steady stress may be caused, at least in part, by the fact that aluminum bronze has a duplex microstructure, whereas monel metal

is a single-phase alloy. The difference, however, may be caused in part by the greater strength of the monel metal and in part by a difference in the permeabilities of the anodic coatings. The influence of stress on the solution pressure would be important only when the stress is locally high enough to cause plastic deformation and when the anodic coating is so permeable that the permeability is not the controlling factor.

As will be shown in a later paper, however, steady stress accelerates the corrosion pitting of steel in distilled water, even when the nominal stress is far below any index of elastic strength. An influence of steady stress on corrosion evidently is favored by conditions causing a slow corrosion rate, with anodic control. Such conditions, when the stress is high enough, tend to cause sharp, deep corrosion pits, which increase the local stress and thus accelerate the advance of the pit.

VIII. SUMMARY

Corrosion pits in the aluminum bronze, unlike the pits in steels, generally are not free to assume a roughly hemispherical form but are confined between crystallites of the alpha phase. Pits in monel metal are more like those in steel. Pits in aluminum bronze and monel metal are much smaller than those in steel. Cyclic stress tends to increase both the size and sharpness of the pits. The effective stress concentration that causes the lowering of the fatigue limit depends on both the size and the sharpness [4]. A sufficiently high combination of stress, cycle frequency, and corrosion time causes fissures to develop from blunt pits. Fissures in the aluminum bronze generally are oblique, whereas fissures in steel and monel metal tend to be perpendicular to the direction of the principal tensile stress.

The curves of decrease of the fatigue limit with corrosion time are of the accelerated-damage type. Such curves are obtained when the corrosion rate is anodically controlled. With these two alloys, therefore, the pits in a specimen have a wider range of size and sharpness than the pits formed under cathodic control, as in steels. The pit that caused the lowering of the fatigue limit may have been much larger and sharper than the pits found in the sectional views. The damage may then be greater than the sectional views would indicate. Many of the sectional views, however, give good correlation with the resultant fatigue limits.

Steady stress tends to accelerate corrosion of aluminum bronze but has little apparent effect on the corrosion pitting of monel metal.

Acknowledgement is made to H. O. Willier for the development of the method of polishing and for the careful preparation and photographing of the metal specimens.

IX. REFERENCES

[1] D. J. McAdam, Jr., *Stress and corrosion*, Report Proc. Third International Congress of Applied Mechanics, Stockholm, **2** 269–293 (1930).

[2] D. J. McAdam, Jr., *Influence of stress on corrosion*, Trans. Am. Inst. Min. Met. Engrs., Inst. of Metals Div. **99**, 282–318 (1932).

[3] D. J. McAdam, Jr., *Stress corrosion of metals*, Proc. Zurich Congress, Int. Assn. Testing Materials **1**, 228–246 (1931).

[4] D. J. McAdam, Jr., and G. W. Geil, *Influence of cyclic stress on the corrosion pitting of steels in fresh water, and the influence of stress corrosion on the fatigue limit*, J. Research NBS **24**, 685 (1940) RP1307.

WASHINGTON, November 29, 1940.

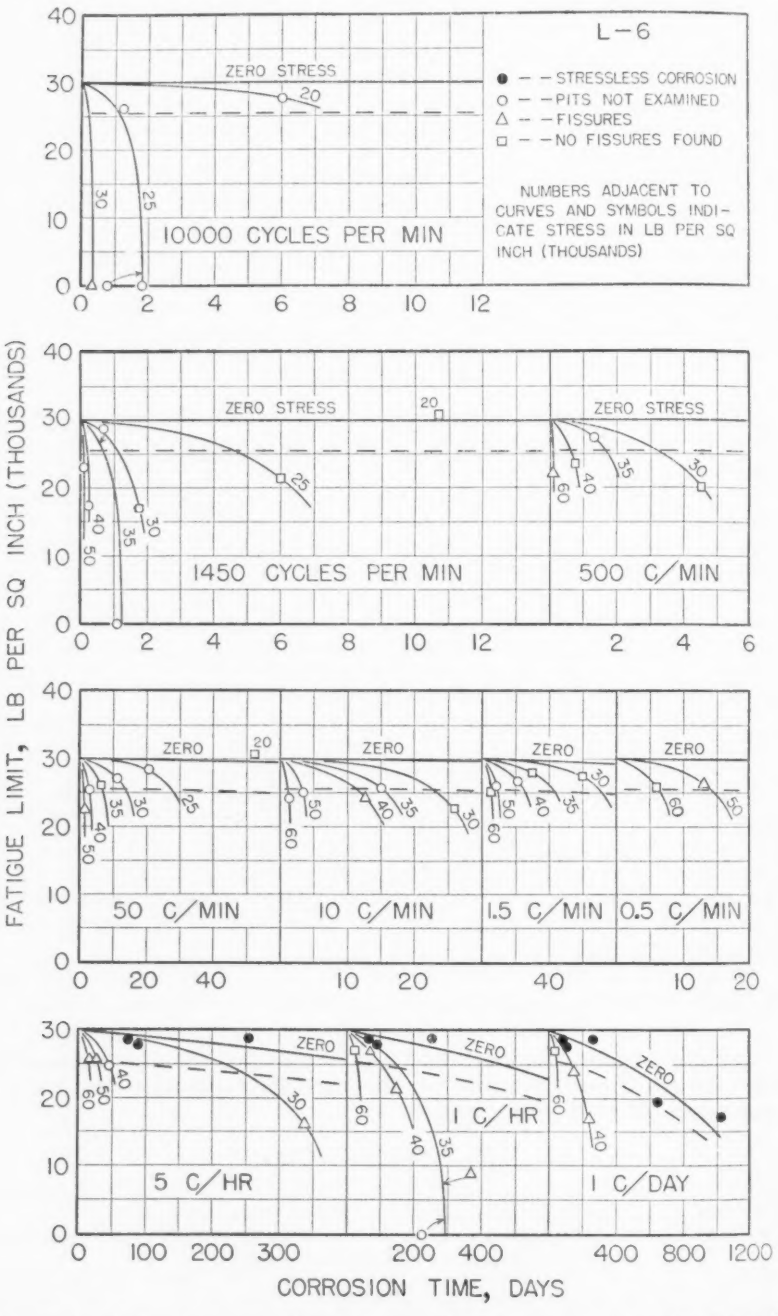


FIGURE 1.—*Influence of corrosion, with and without stress, on the fatigue limit of aluminum bronze.*

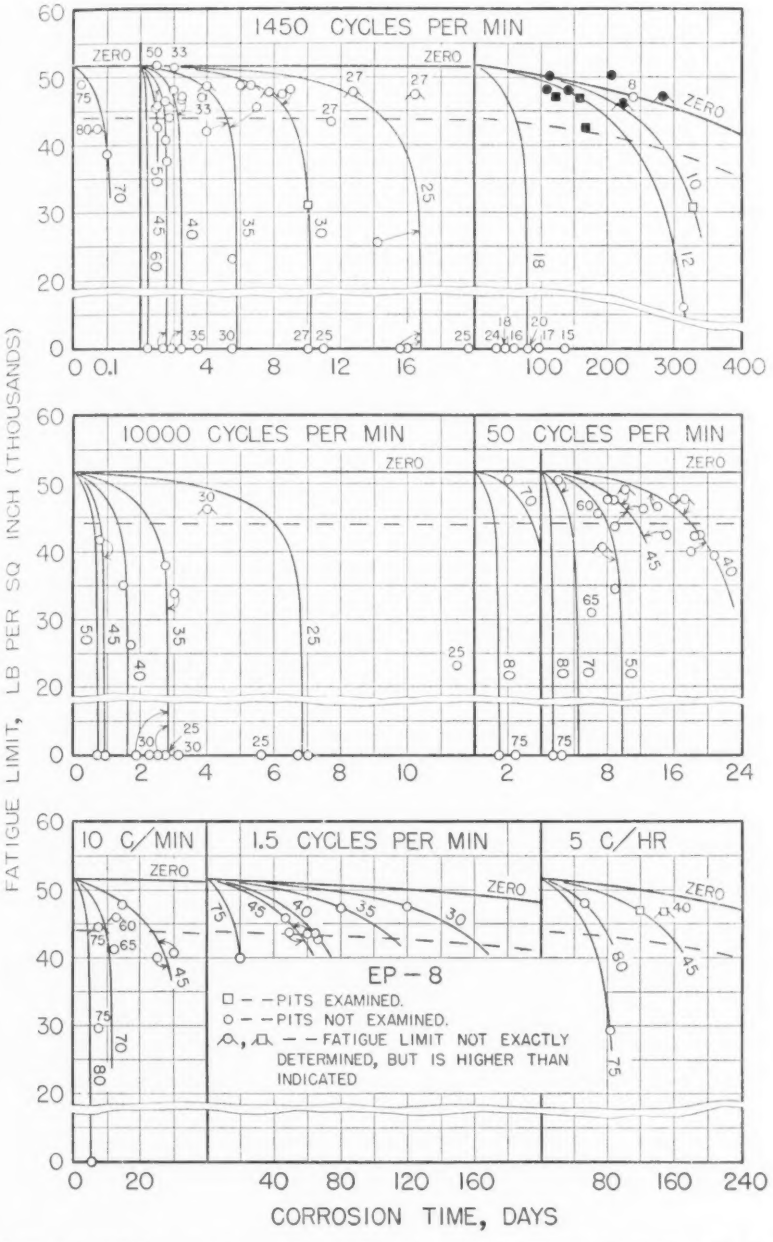

1450 CYCLES PER MIN
10000 CYCLES PER MIN
50 CYCLES PER MIN
10 C/MIN
1.5 CYCLES PER MIN
5 C/HR
ZERO
FATIGUE LIMIT, LB PER SQ INCH (THOUSANDS)
CORROSION TIME, DAYS
EP-8
□ - - PITS EXAMINED.
○ - - PITS NOT EXAMINED.
△, ▽ - - FATIGUE LIMIT NOT EXACTLY DETERMINED, BUT IS HIGHER THAN INDICATED

FIGURE 2.—*Influence of corrosion, with and without stress, on the fatigue limit of monel metal.*

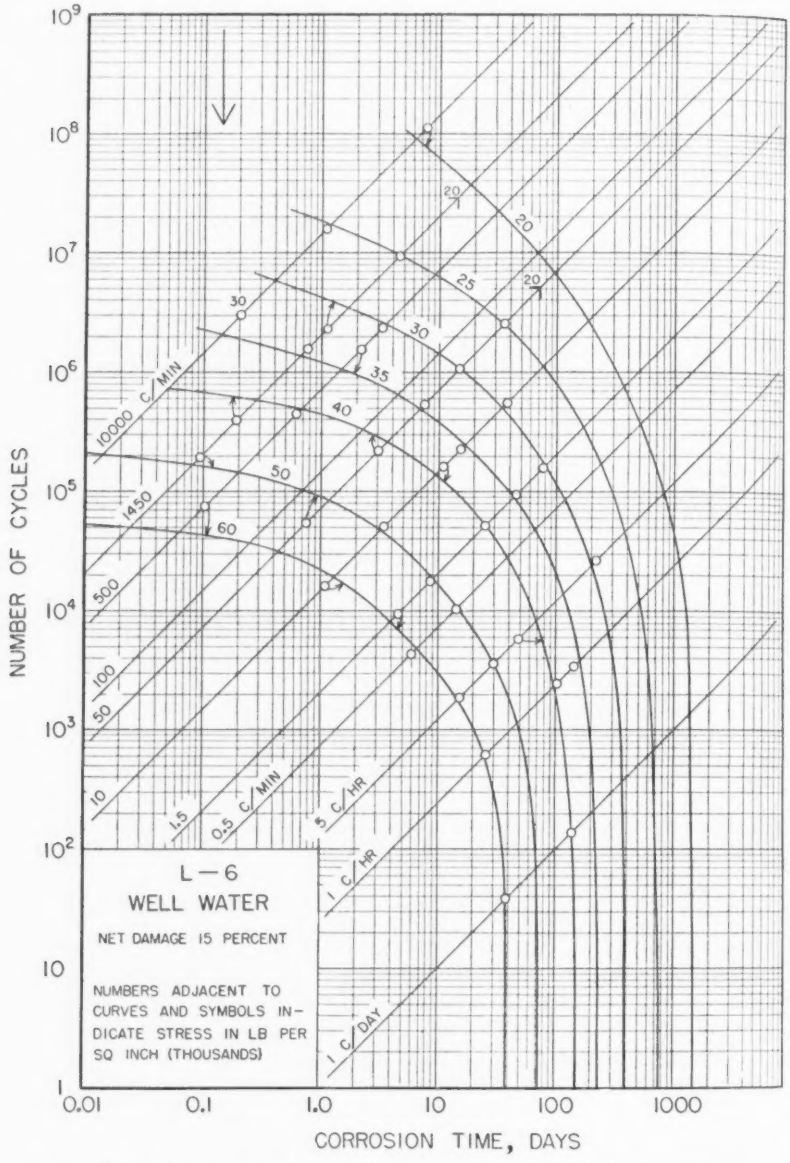


FIGURE 3.—*Constant net damage due to stress corrosion of aluminum bronze.*

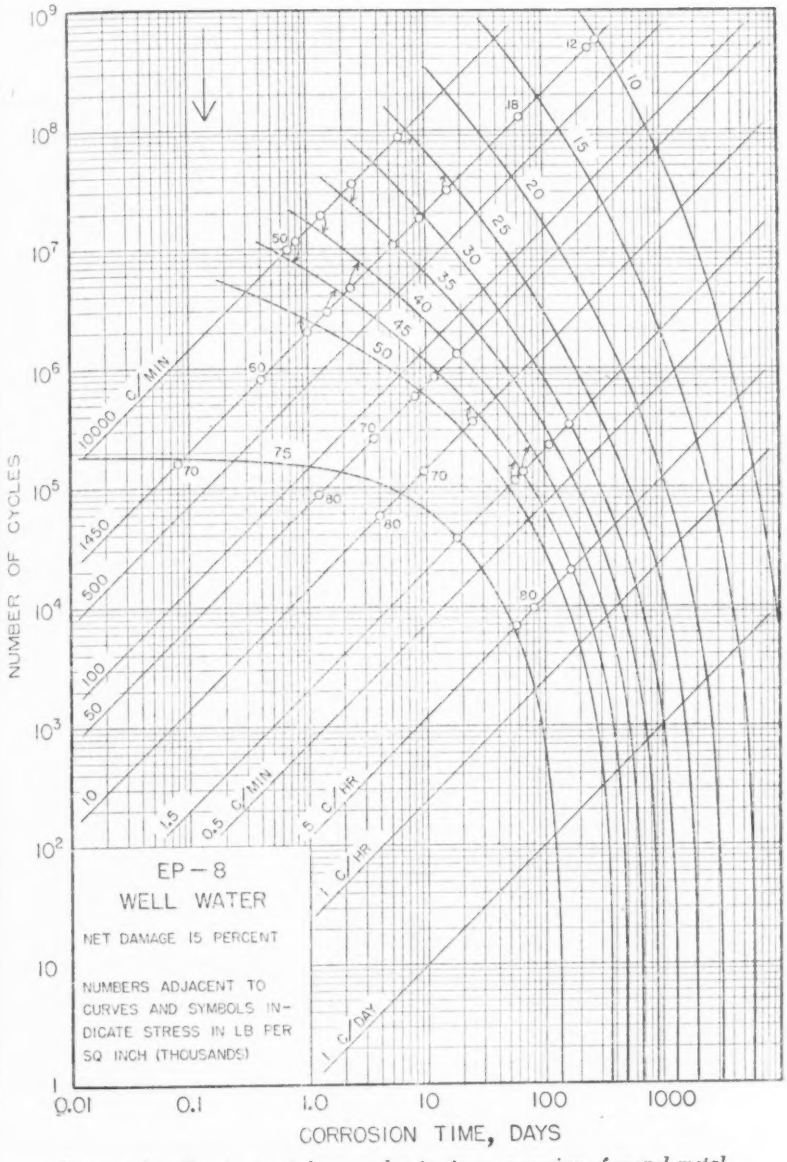


FIGURE 4.—*Constant net damage due to stress corrosion of monel metal.*

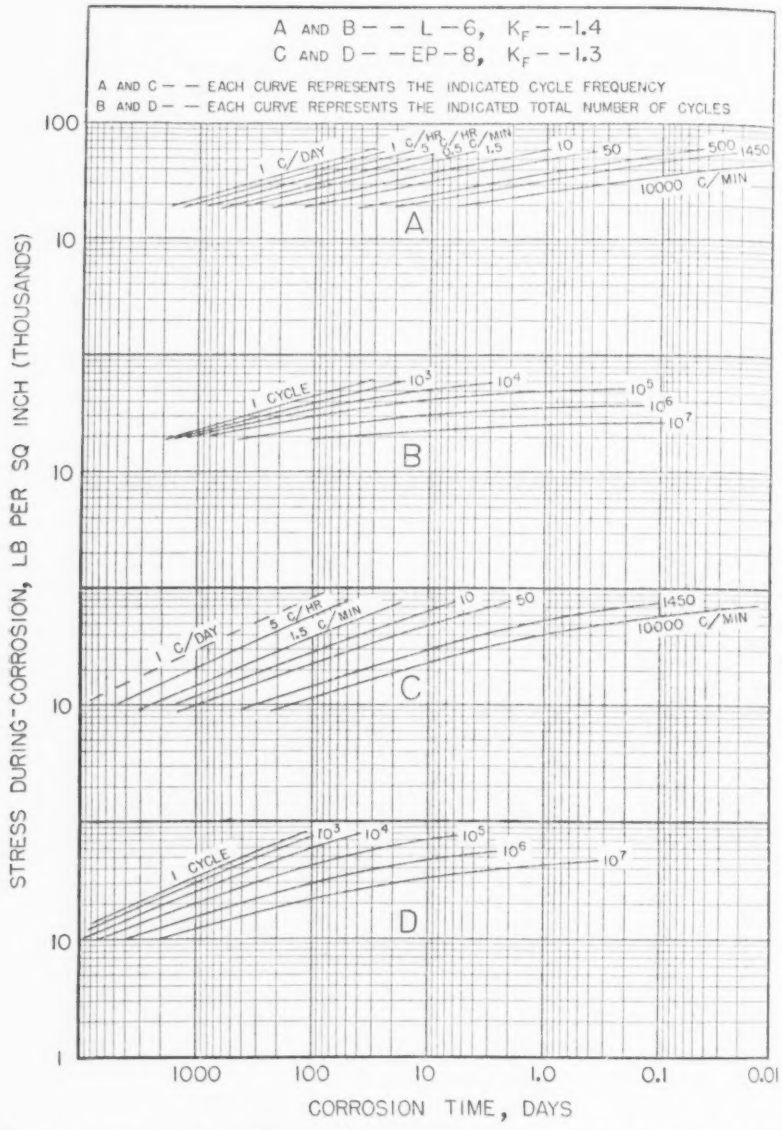


FIGURE 5.—*Constant net damage due to stress corrosion of aluminum bronze and monel metal.*

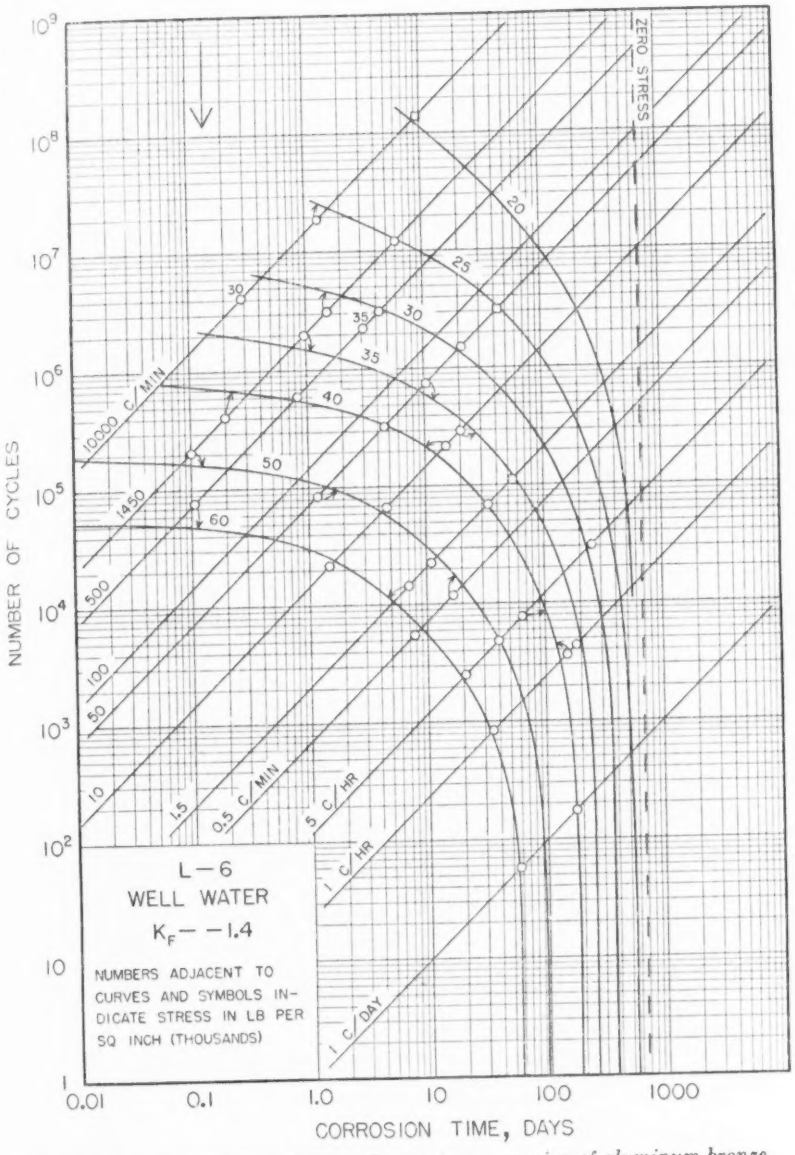


FIGURE 6.—*Constant total damage due to stress corrosion of aluminum bronze.*

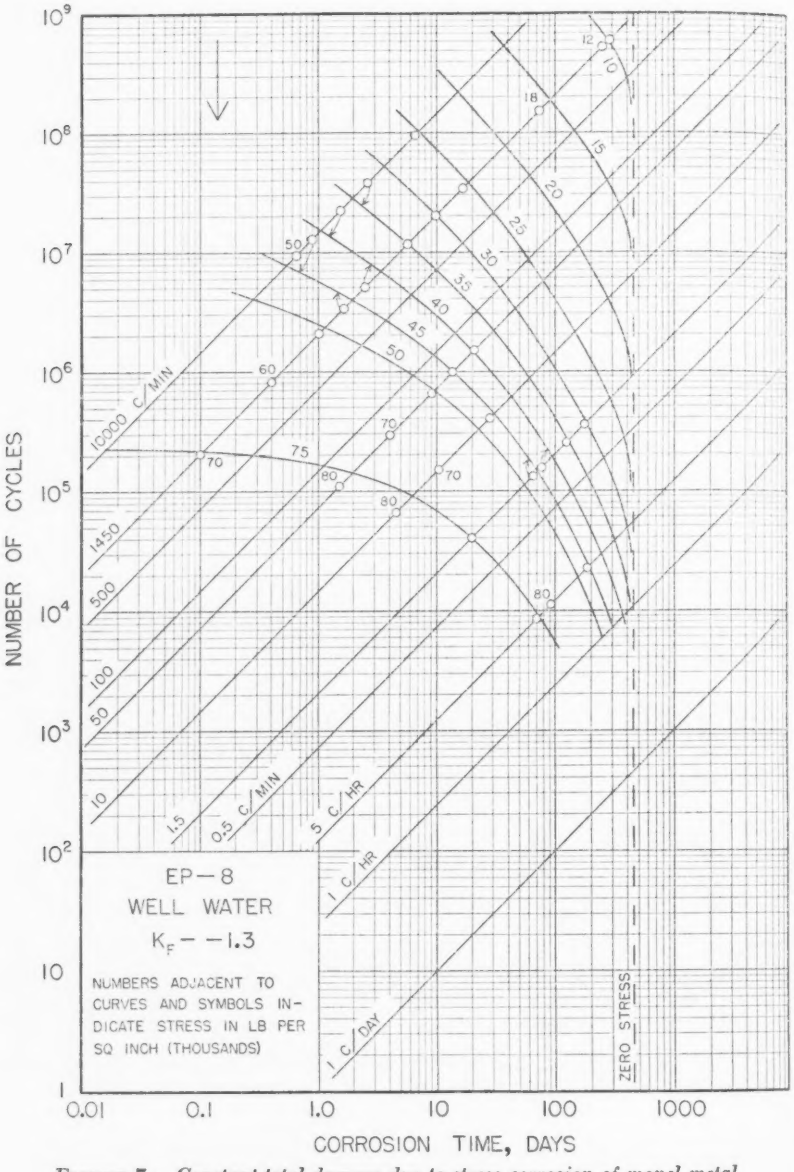


FIGURE 7.—*Constant total damage due to stress corrosion of monel metal.*

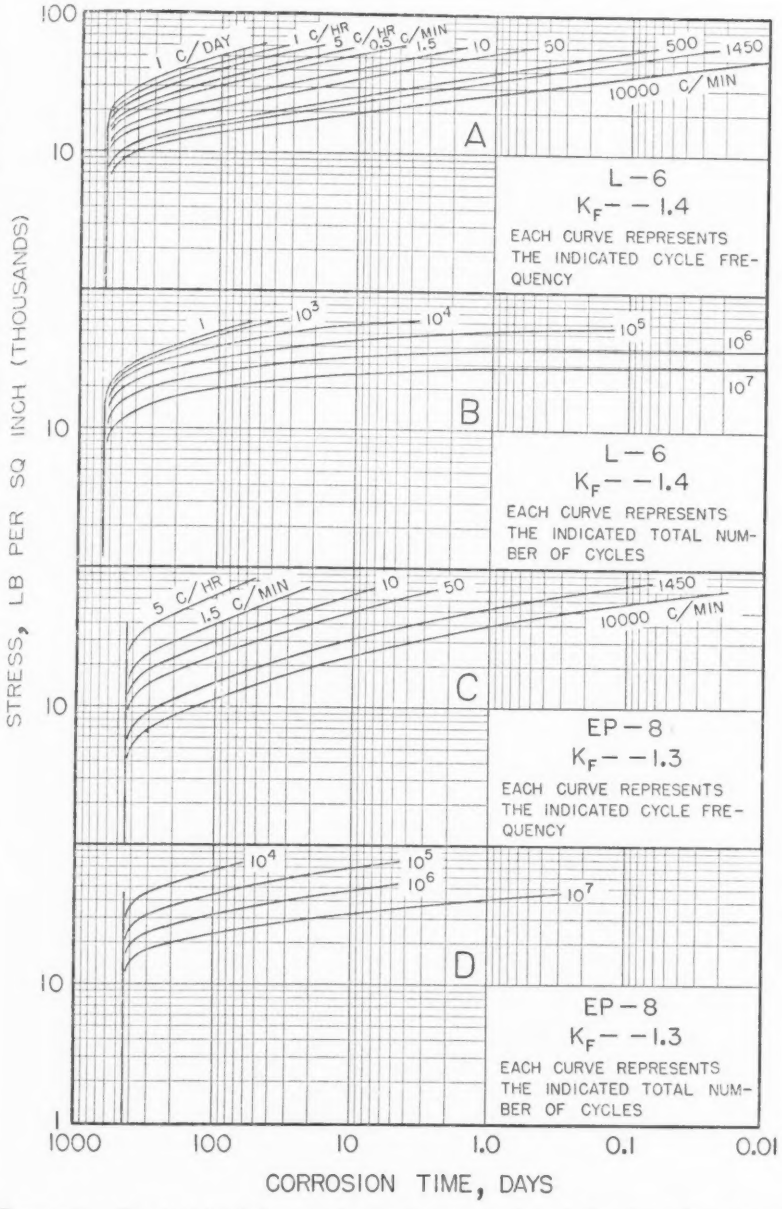


FIGURE 8.—*Constant total damage due to stress corrosion of aluminum bronze and monel metal.*

UNIV. OF MICH.

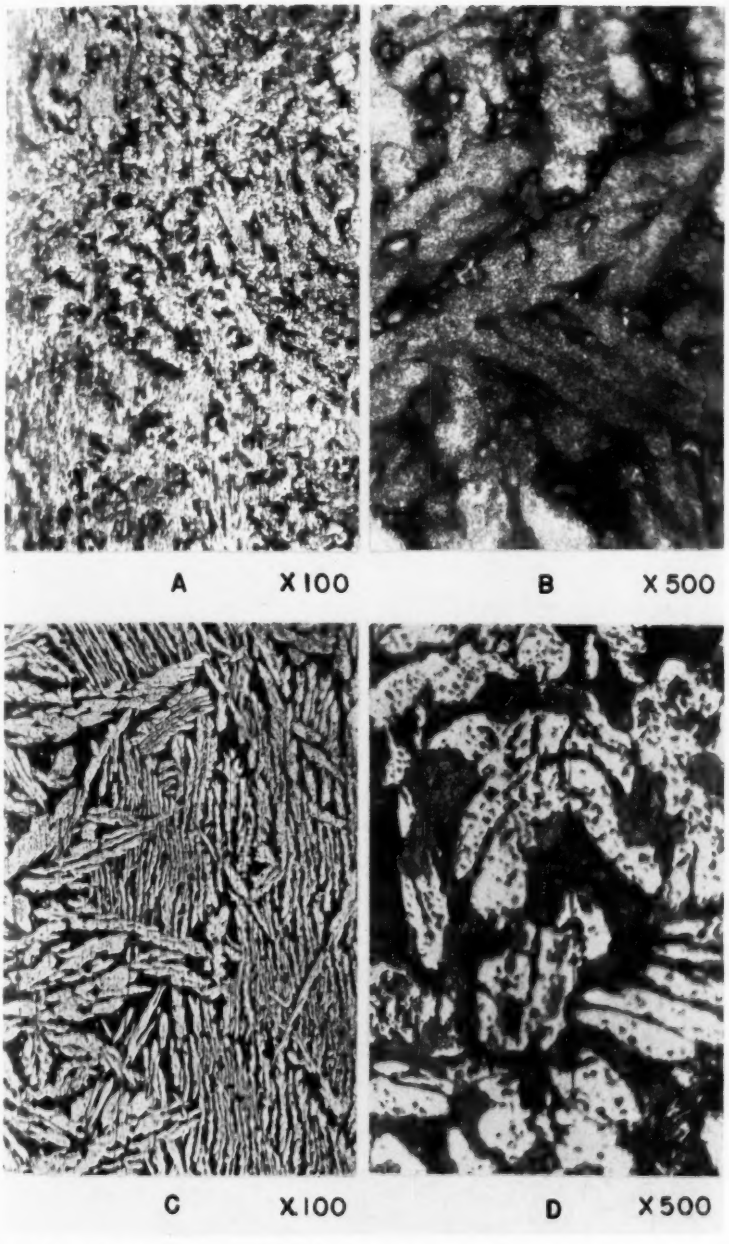


FIGURE 9.—*Aluminum bronze, 1 cycle per day, surface views, Severn River water, ×4.*

60,000 lb/in.², 16 days.

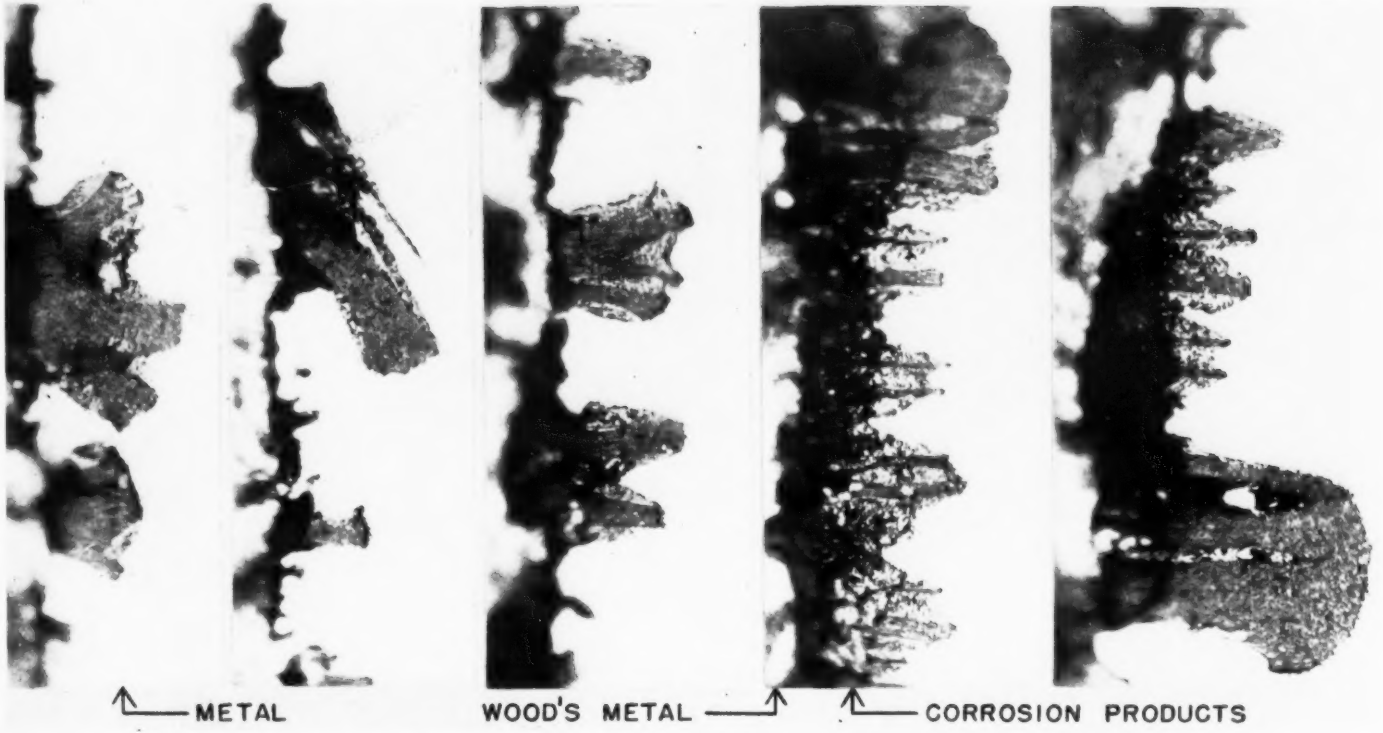


FIGURE 10.—*Aluminum bronze, 1 cycle per day, longitudinal sections, Severn River water, unetched,* ×1000.

(91,000 Dec. 7, 16 days.)

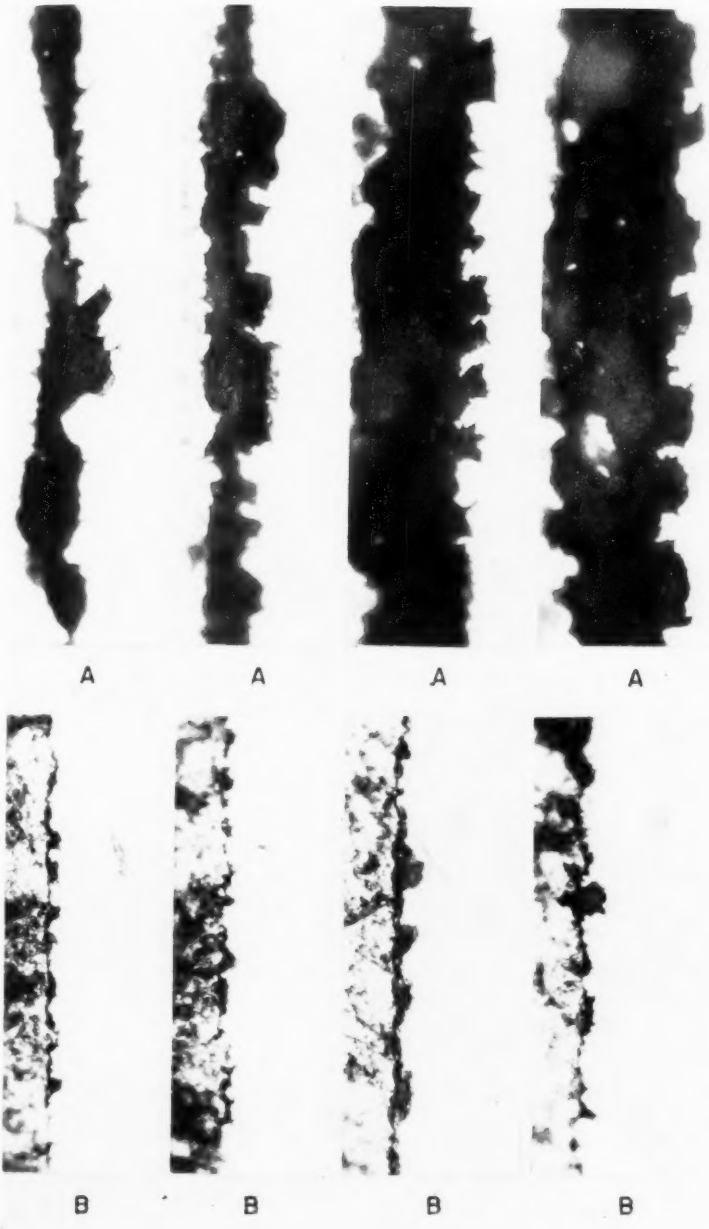


FIGURE 11.—*Aluminum bronze, stressless corrosion, longitudinal sections, well water ×1000.*

A, 66 days.
B, 88 days.

www.nbs.gov/jres

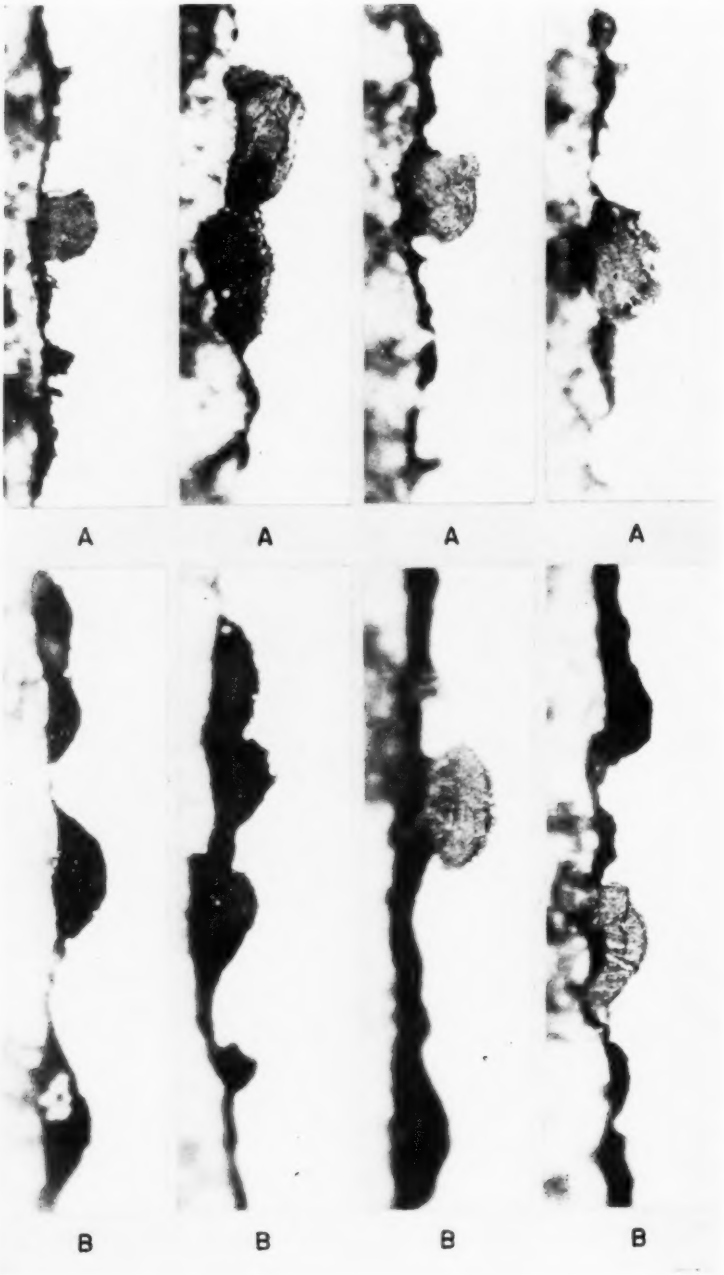


FIGURE 12.—*Aluminum bronze, stressless corrosion, longitudinal sections, well water, ×1000.*

A, 640 days.
B, 1,021 days.

FIGURE 13.—*Aluminum bronze, 1,450 cycles per minute, longitudinal sections, well water, ×1000.*

A, 30,000 lb/in.², 1.8 days.
B, 25,000 lb/in.², 6 days.
C, 20,000 lb/in.², 11 days.

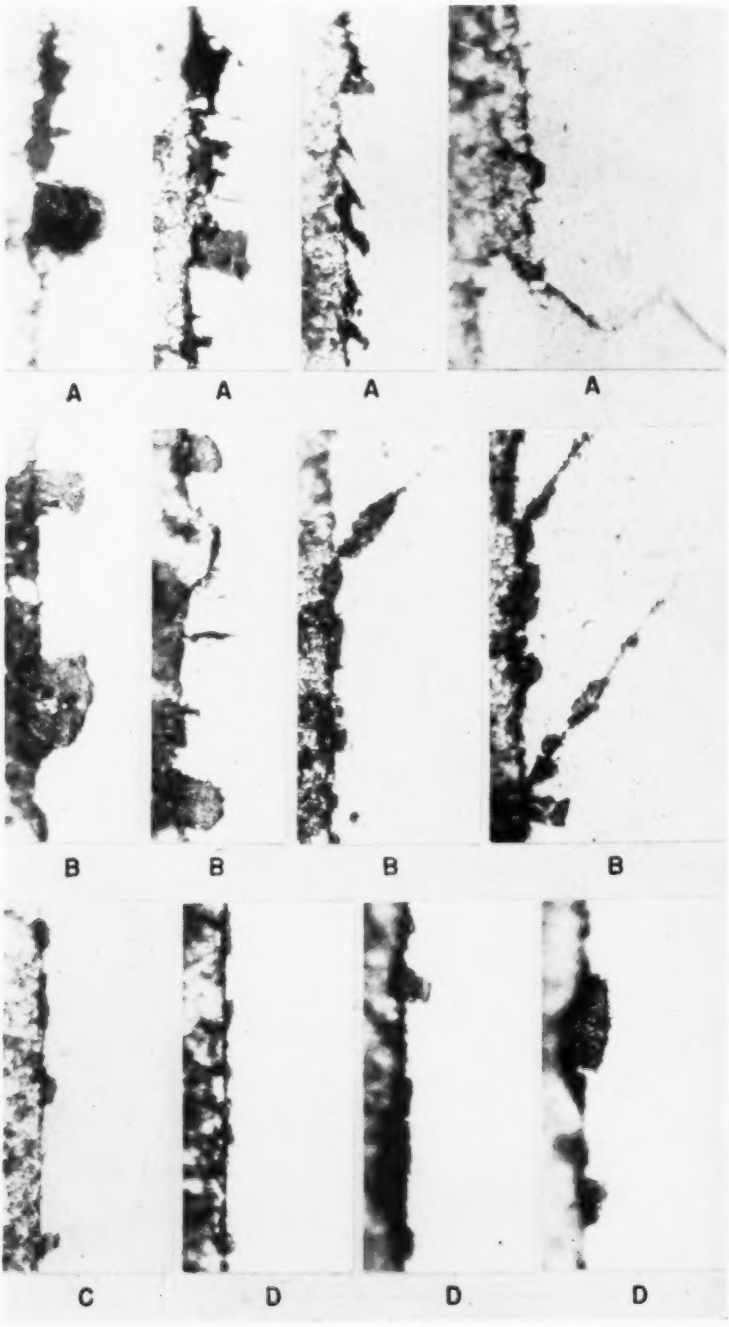


FIGURE 14.—*Aluminum bronze, longitudinal sections, well water, ×1000.*

A, 10,000 cycles per minute, 30,000 lb/in.^{2}, 0.3 day.
B, 500 cycles per minute, 60,000 lb/in.^{2}, 0.06 day.
C, 500 cycles per minute, 40,000 lb/in.^{2}, 0.7 day.
D, 500 cycles per minute, 30,000 lb/in.^{2}, 4.5 days.

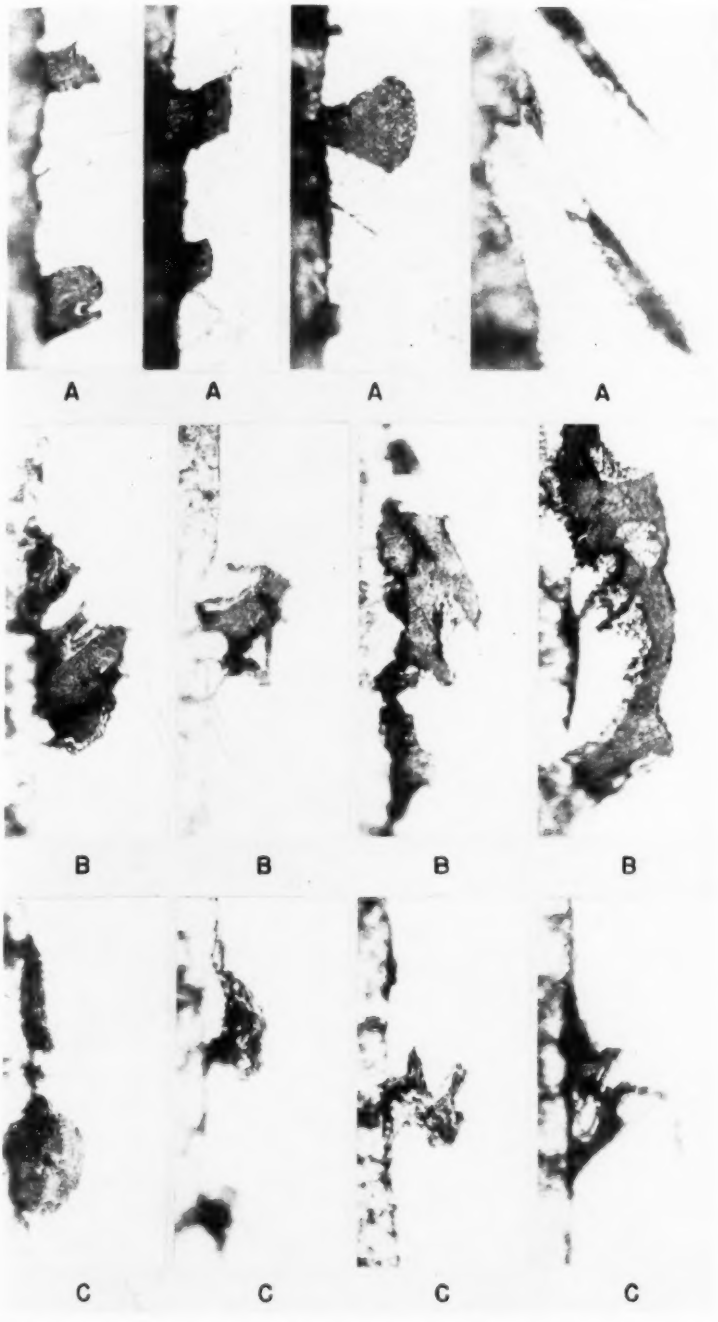


FIGURE 15.—*Aluminum bronze, 50 cycles per minute, longitudinal sections, well water, ×1000.*

A, 50,000 lb/in.², 1.1 days.
B, 35,000 lb/in.², 6.7 days.
C, 20,000 lb/in.², 52 days.

FIGURE 16.—*Aluminum bronze, 10 cycles per minute, longitudinal sections, well water, ×1000.*

A, 40,000 lb/in.², 13 days.
B, 30,000 lb/in.², 26 days.

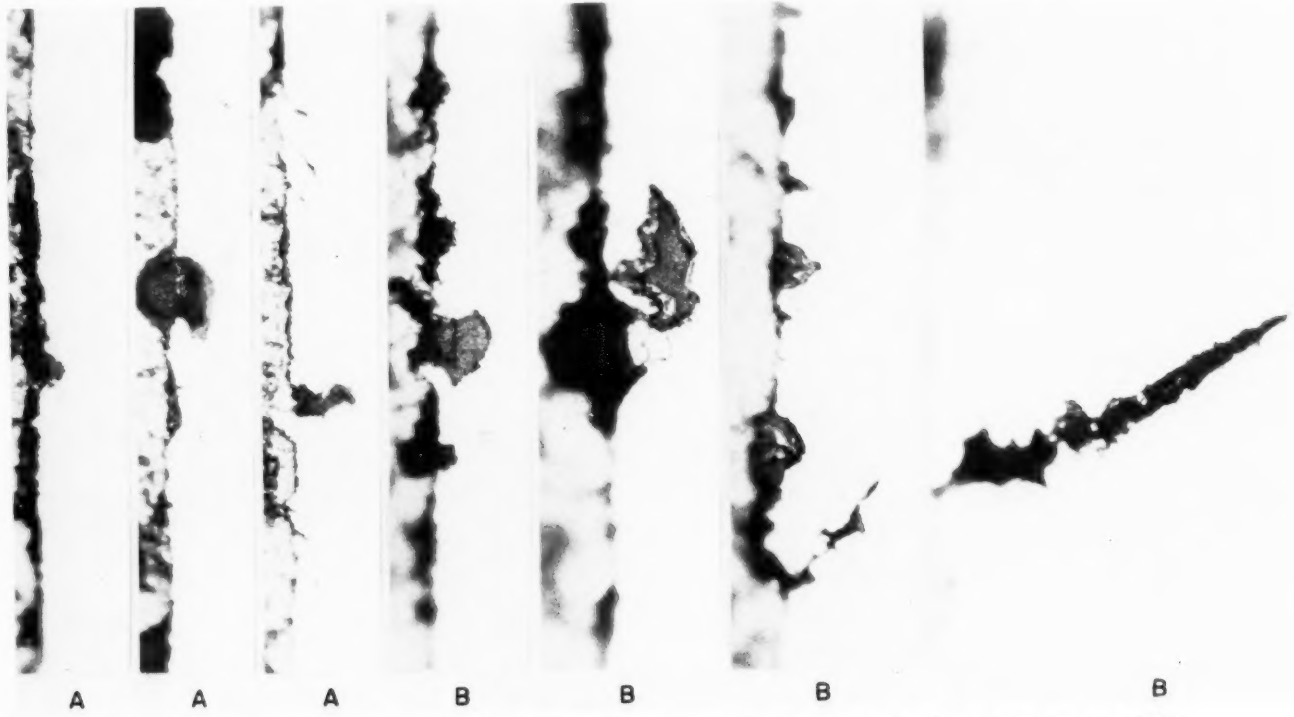


FIGURE 17.—*Aluminum bronze, 0.5 cycle per minute, longitudinal sections, well water, ×1000.*

A, 60,000 lb/in.², 6 days.
B, 50,000 lb/in.², 13 days.

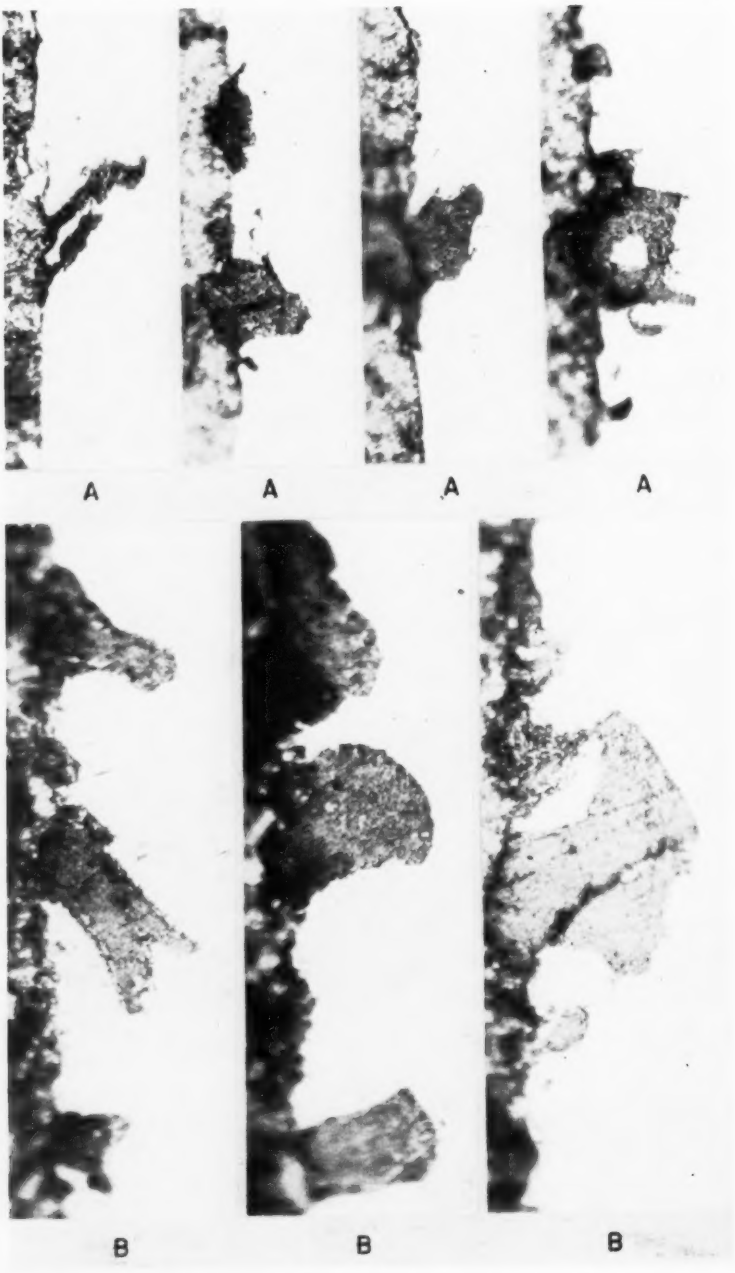


FIGURE 18.—*Aluminum bronze, 5 cycles per hour, longitudinal sections, well water, ×1000.*

A, 60,000 lb/in.², 15 days.
B, 50,000 lb/in.², 24 days.

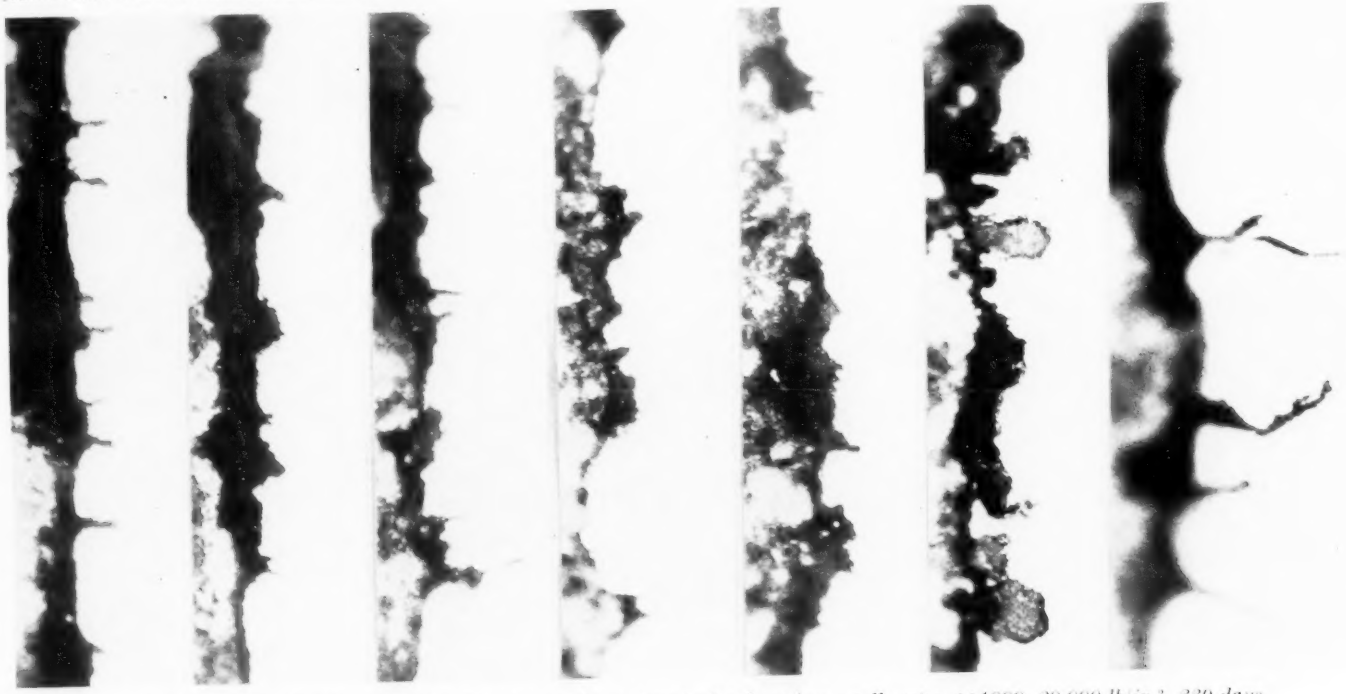


FIGURE 19.—*Aluminum bronze, 5 cycles per hour, longitudinal sections, well water, ×1000, 30,000 lb/in.², 339 days.*

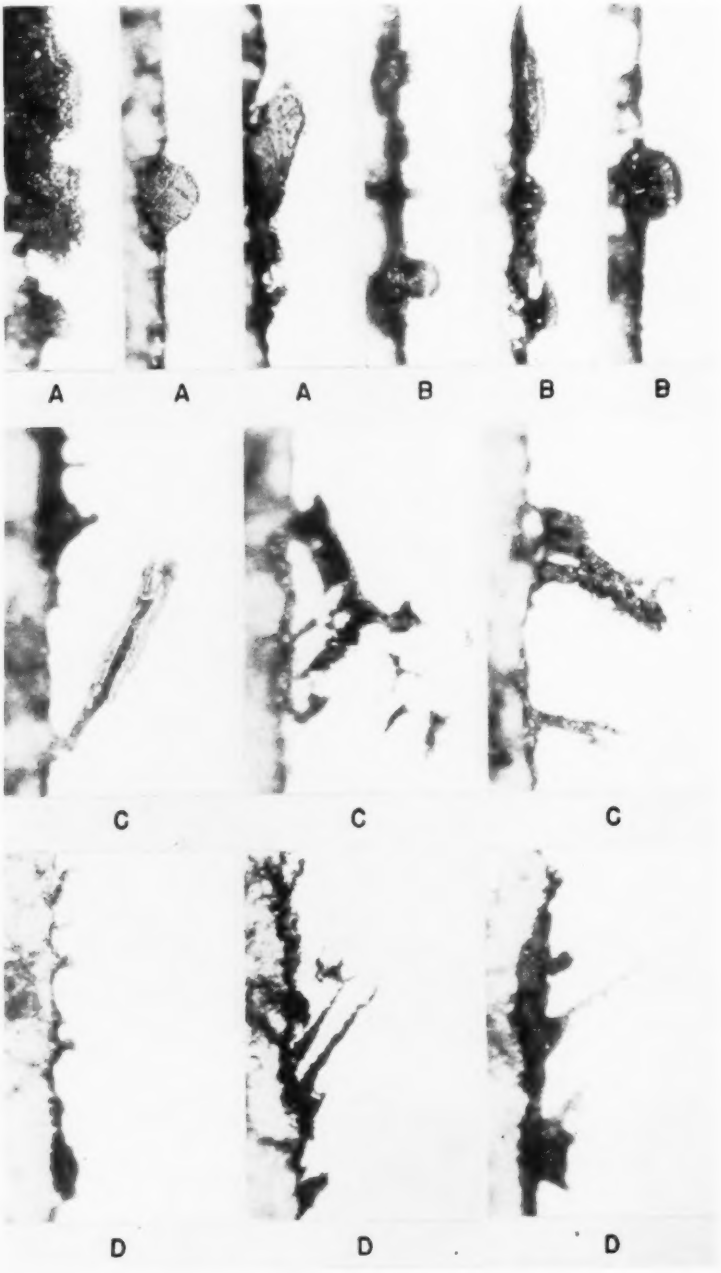


FIGURE 20.—*Aluminum bronze, 1 cycle per hour, longitudinal sections, well water, ×1000.*

A, 60,000 lb/in.², 20 days.
B, 40,000 lb/in.², 73 days.
C, 40,000 lb/in.², 149 days.
D, 35,000 lb/in.², 374 days.

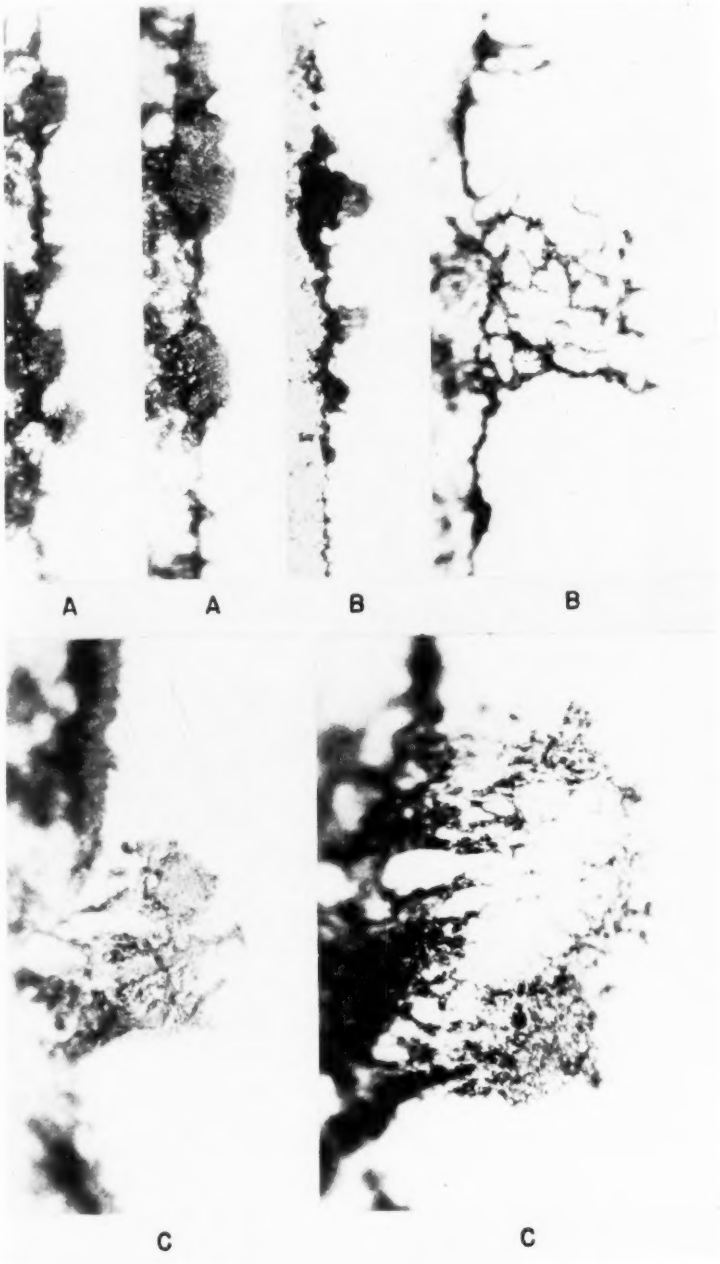


FIGURE 21.—*Aluminum bronze, 1 cycle per day, longitudinal sections, well water, ×1000.*

A, 60,000 lb/in.², 28 days.
B, 40,000 lb/in.², 151 days.
C, 40,000 lb/in.², 240 days.

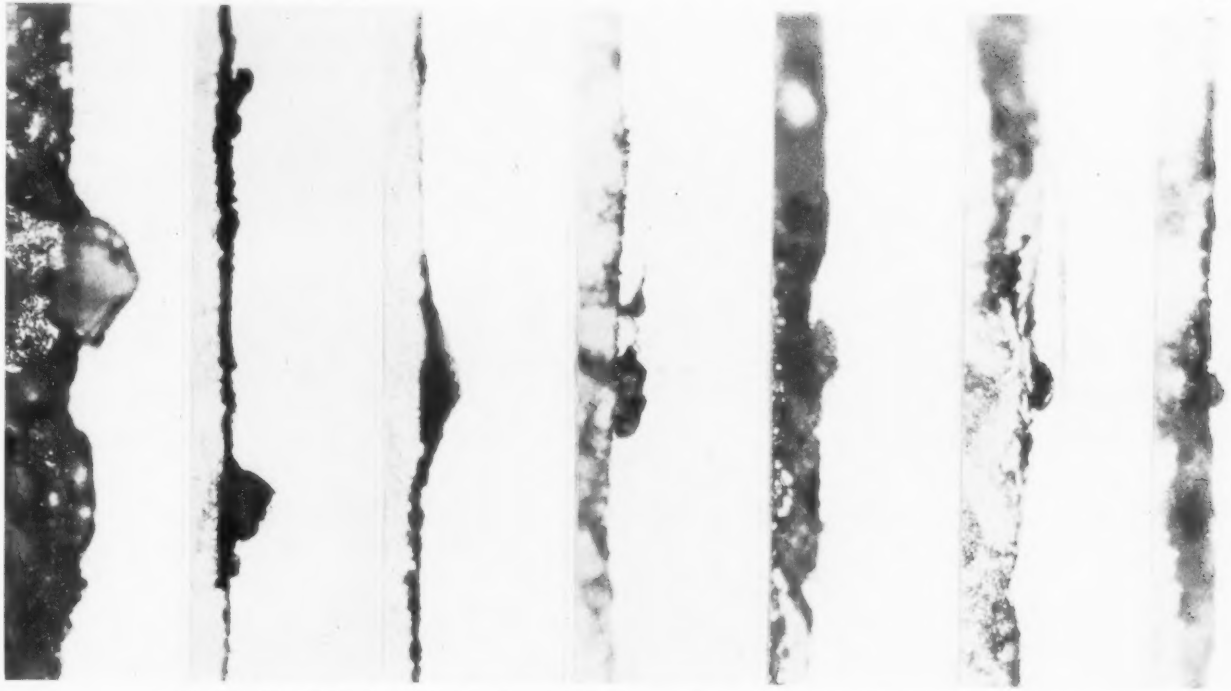


FIGURE 22.—*Monel metal, stressless corrosion, longitudinal sections, well water, ×1000. 122 days.*

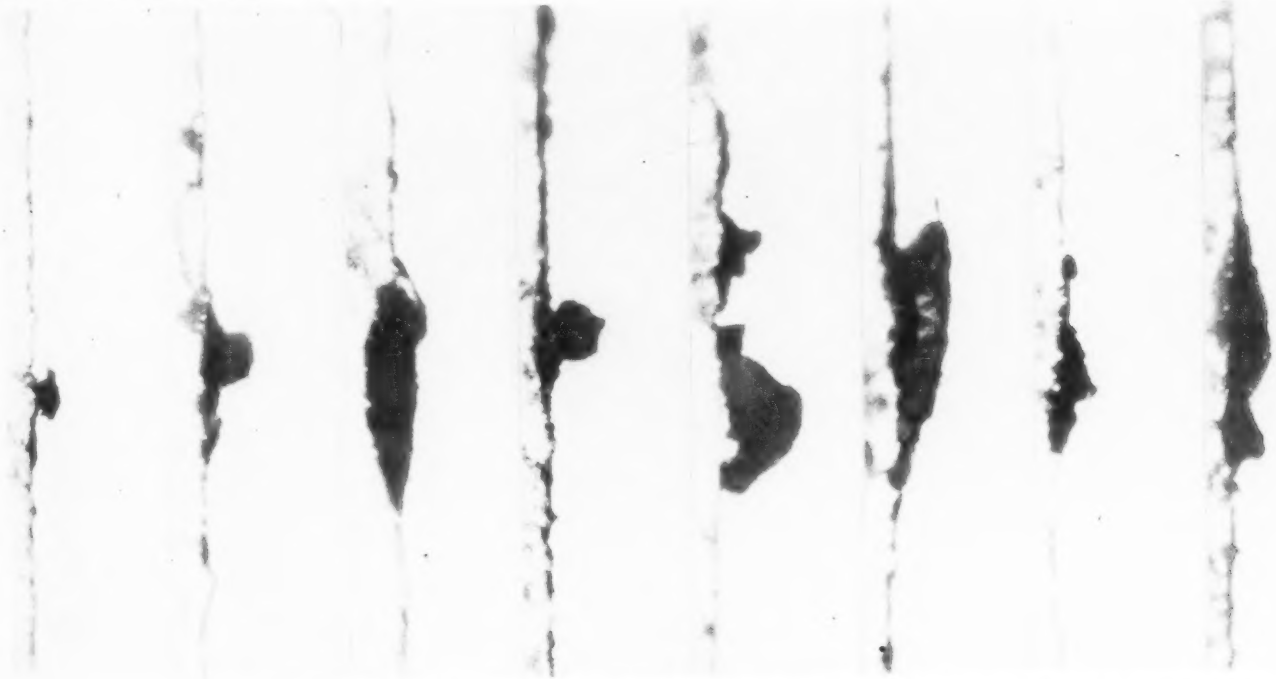


FIGURE 23.—*Monel metal, stressless corrosion, longitudinal sections, well water, ×1000, 157 days.*

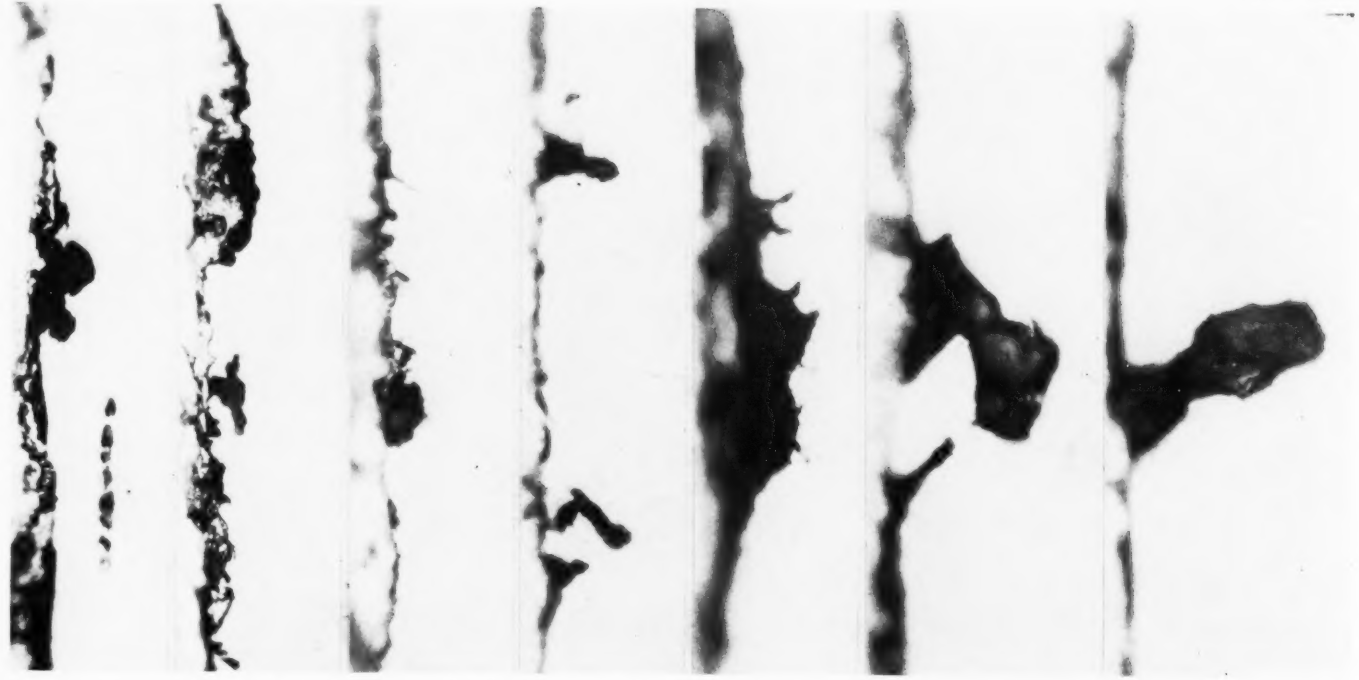


FIGURE 24.—*Monel metal, stressless corrosion, longitudinal sections, well water, ×1000. 170 days.*

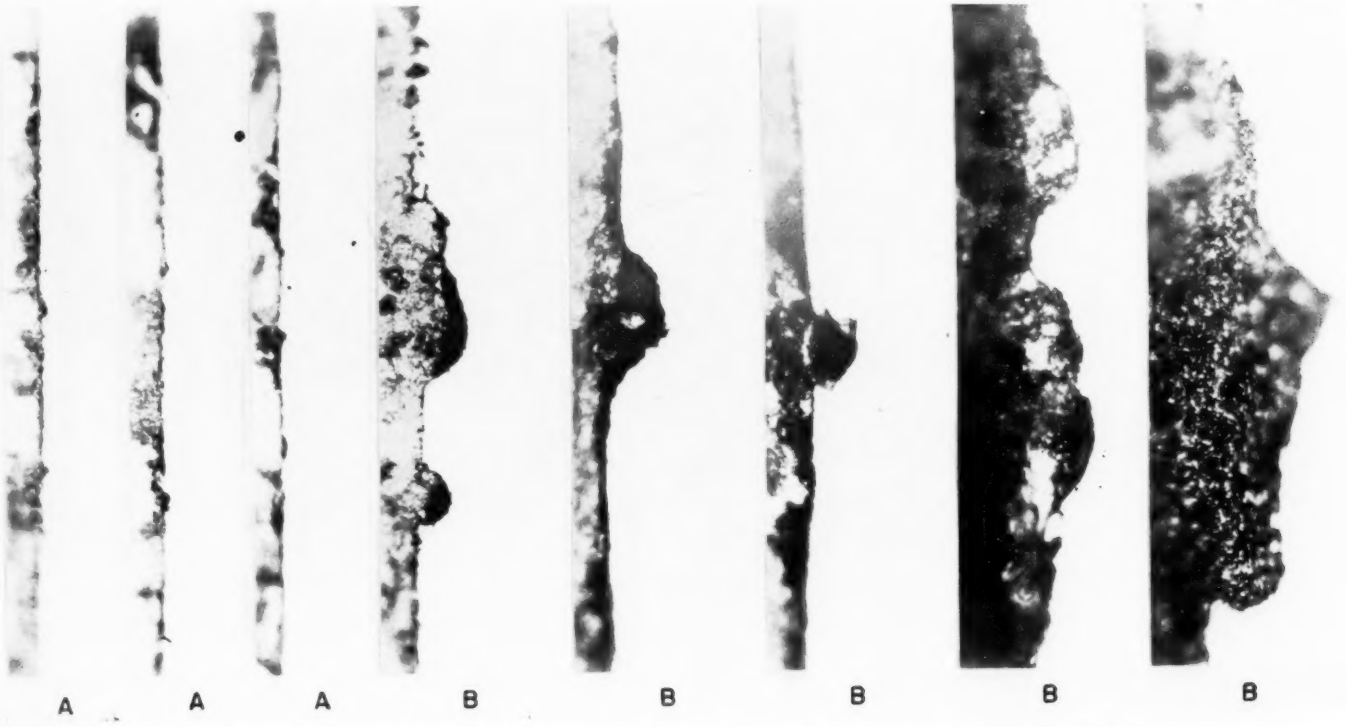


FIGURE 25.—*Monel metal, 1,450 cycles per minute, longitudinal sections, well water, ×1000.*

A, 30,000 lb/in.², 10 days.
B, 10,000 lb/in.², 331 days.

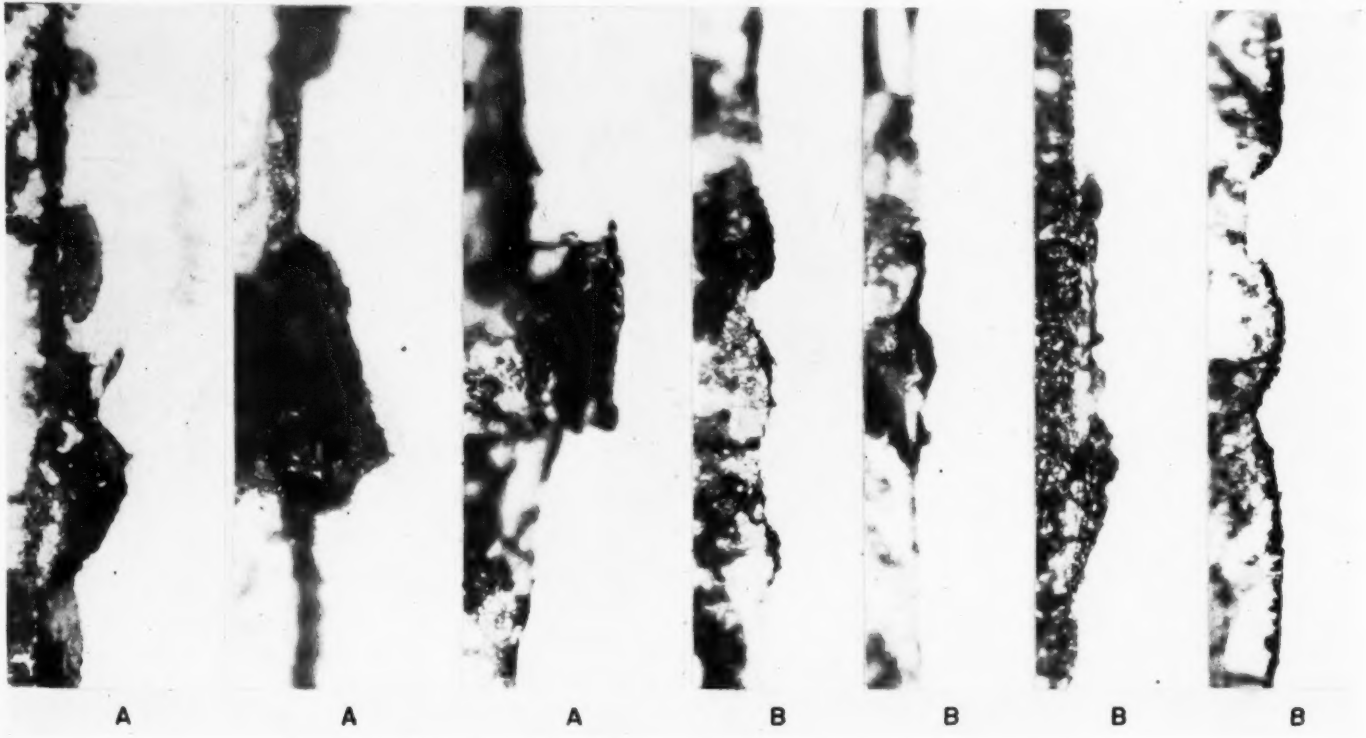


FIGURE 26.—*Monel metal, 5 cycles per hour, longitudinal sections, well water,* $\times 1000$.

A, 45,000 lb/in.^{2}, 123 days.
B, 30,000 lb/in.^{2}, 148 days.

U. S. DEPARTMENT OF COMMERCE NATIONAL BUREAU OF STANDARDS

RESEARCH PAPER RP1367

Part of Journal of Research of the National Bureau of Standards, Volume 26, February 1941

DISTRIBUTION OF OZONE IN THE STRATOSPHERE: MEASUREMENTS OF 1939 AND 1940

By W. W. Coblentz and R. Stair

ABSTRACT

Continuing the measurements of ultraviolet solar radiation intensities in the stratosphere made in 1937 and 1938 (RP1207), in this paper a description is given of similar measurements made with improved apparatus in the summer of 1939 and 1940.

The apparatus that was transported aloft by means of unmanned balloons consisted of the following elements: (1) a photoelectric cell and filter radiometer, connected with a balanced amplifier, relaxation oscillator, and radio transmitter for measuring the ultraviolet intensities; (2) a radio-operated barometer, giving the heights attained by the apparatus; and (3) an electrolytic resistor for measuring the ambient temperature of the ultraviolet intensity meter.

The audio-frequency modulated radio wave, giving (1) the intensity of the incident ultraviolet solar radiation, (2) the elevation of the apparatus, and (3) the temperature of the ultraviolet meter, was received and graphically recorded at a fixed ground station (RP1075).

In three successful flights made in June 1939, and in two similar flights made in 1940, the ultraviolet radiometer attained heights ranging from about 78,000 to about 88,000 ft (23 to 27 km). From the observed change in filter transmissions of ultraviolet solar radiation with elevation, it is deduced that the apparatus passed through 65 to 70 percent of the ozone (and other ultraviolet-absorbing constituents, if any, in the upper atmosphere), the most of which is localized in a layer extending from 15 to 27 km, with a wide maximum of concentration at a height of about 24 to 25 km, above the earth's surface. At the highest elevation attained by the instruments, the intensity of the ultraviolet solar radiation of wavelengths shorter than 3132 A was 10 to 12 times that observed at sea level.

CONTENTS

	Page
I. Introduction	161
1. Tests with the new type radio tubes	163
II. Exploration of the vertical distribution of ozone in the stratosphere	163
1. Notes on the stratosphere balloon flights	164
III. Evaluation of the ultraviolet intensity measurements	165
1. Effect of temperature on the ultraviolet measurements	166
2. The extraterrestrial ultraviolet solar spectral-energy curve	167
3. Exposition of data	169
IV. References and notes	173

I. INTRODUCTION

In a previous communication [1][1] a description was given of the results obtained in June 1938 in measuring ultraviolet solar intensities, and thereby determining the vertical distribution of atmospheric ozone, at various elevations above the earth's surface. Readers interested in the subject are referred to this and other papers for a description of the instruments used [1, 6, 7]; the methods of measure-

[1] Figures in brackets indicate the literature references and notes at the end of this paper.

ment, and reduction of observations; and for a discussion of the preliminary measurements made in June 1937 [5, 6].

For clarity in exposition of the herein described data, it is relevant to record that an ultraviolet-intensity meter, consisting of a photoelectric cell and filter radiometer connected with an audio-frequency generator and radio transmitter, is transported aloft by means of small unmanned balloons [5]. The radio signals giving the altitude and temperature of the apparatus, and the ultraviolet solar intensities, are received and recorded graphically at a fixed ground station [5, 6, 7].

As mentioned in the previous paper [1], the problem then was (and still is) the lack of a satisfactory photoelectric cell for measuring the incident solar ultraviolet radiation, which increases rapidly in intensity and changes markedly in spectral quality, as the apparatus rises above the earth's surface and passes through the layer of ozone that is concentrated principally in the stratosphere.

The type of photoelectric cell now available commercially, best suited for this work, consists of a spherical bulb covered with a diffusing window; the interior of the bulb being photoelectrically sensitized with a suitable metal (Cd) which, in the samples used, changed in spectral response and intrinsic sensitivity with age, and with temperature, to a greater extent than is desirable.[2]

One of these photoelectric cells (Cd–38303) had a unique record of performance. It made three stratosphere flights (1938, 1939, and 1940), and was used in measuring sky and solar radiation at Flagstaff, Ariz., in 1938. In its flight of June 20, 1940, after attaining an altitude of 83,000 ft (25 km), the radio sonde descended to an elevation of 70,000 ft (21 km), where it floated for several hours and was carried out to sea. The following day, at sunset, it was seen falling into the sea with three of the original six balloons still inflated, some 140 miles southeast of the Boston Lightship, and was recovered soon thereafter. The report received with the return of the instrument, a month later, stated: "On delivery at Fish Pier, Boston, box was badly broken, contents loosened from fastenings; but (radio) tubes not broken. Batteries, (barograph, motor, filters) were missing."

It is interesting to note that the spectral response of Cd–38303 (as indicated by the transmissions of four filters [17], using a Mazda *CX* lamp as a source of ultraviolet) before the flight of 1940 was the same as before the flight of 1939. Similarly, the spectral response after the flight of 1940 was the same as after the flight made in 1939. In both years the percentage transmissions of ultraviolet through these filters was lower (indicating a higher spectral response in the short wavelengths relative to the long wavelengths) after the flight than before the flight, with an apparent recovery in relative spectral response after resting a year. Similar changes in spectral response were observed in other cadmium-photoelectric cells after a flight. A correction for this change in filter transmission, which is similar in magnitude but opposite in direction to the correction for temperature, has not been applied to the herein described ultraviolet measurements, because it is small and because of the lack of knowledge as to when the change occurred during a flight.

[2] In this connection, we record our appreciation of the assistance of Dr. Louis Koller, of the Research Laboratory of the General Electric Company, in selecting photoelectric cells that had closely the spectral range of sensitivity desired in this investigation. We thank also the National Geographic Society and the Radio Section of the National Bureau of Standards for their cooperation in this work.

1. TESTS WITH THE NEW TYPE RADIO TUBES

In view of the fact that the highest elevation attainable is limited by the kind of balloons used and by the weight of the apparatus that is being transported, preliminary to the flights of 1940 an attempt was made to reduce the weight by employing smaller batteries and radio tubes than were previously used in the ultraviolet meter [1, 5, 7].

At the outset it was found that the small radio tubes, recently introduced for use in hearing aids, were unsuitable for operation in the stratosphere ultraviolet meter.

Several radio units were assembled using some of the new 1.4-volt radio tubes (principally the types 1N5-GT, 1A7-GT, and 1G6-G), in an attempt to decrease the weight of the apparatus. However, since the battery requirements were similar to those used with the older type of radio tubes, the actual saving in weight in the ultraviolet meter was only about 4 ounces.

The 1N5-GT tubes operated satisfactorily, and were incorporated in some of the amplifiers used in 1940. However, none of the combinations of the other new models of radio tubes investigated operated as satisfactorily as those previously employed in the transmitters, and their use without further prolonged experiments was considered inexpedient, especially since there would have been but little reduction in weight of the radio sonde.

II. EXPLORATION OF THE VERTICAL DISTRIBUTION OF OZONE IN THE STRATOSPHERE

Reference has already been made to earlier papers [1, 5, 6, 7] for a description of apparatus and methods of measurement of ultraviolet solar intensities in the stratosphere.

The balloon flights of 1939 differed from those made in 1940 only in that, in the latter, additional equipment in the form of an electrolytic resistor [2, 3] was included to measure the ambient temperature of the photoelectric cell and filters. This temperature element was near the center of the axis of rotation of the disk that supported the filters, and hence measured the temperature of the enclosure under the disk and the cover of the box containing the ultraviolet meter (see fig. 6 of RP1207 [1]). As in previous years, the top of the box ("cover") was painted black to heat the interior by absorbing solar radiation. The sides and bottom of the box were wrapped in cotton batting and black cloth to retard cooling of the batteries. An extension arm, attached to the axis of rotation of the disk that supports the filters, provided a means for opening the electric circuit of the ultraviolet meter and closing the circuit through the temperature element. In order to eliminate interference between the radio signals indicating the ultraviolet intensities, and the temperature signal, all the filters and screens were placed close together on one side of the disk, leaving about one-half the disk opaque (see fig. 6, of RP1207 [1]). As this opaque part of the disk passed over the window of the photoelectric cell, two records per revolution of the disk were obtained of the reference frequency of the ultraviolet meter (see figs. 8 and 9, RP1207 [1]); and when the temperature element was in the circuit, two intervening frequency measurements were recorded, giving the temperature. Since the disk made one revolution in about 45 seconds, and the rate of ascent was 700 to

1,000 ft/min, numerous measurements were obtained of the ultraviolet solar intensity and of the temperature of the enclosure at various elevations above the earth's surface.

1. NOTES ON THE STRATOSPHERE BALLOON FLIGHTS

In the flights of 1939, the weight of the complete apparatus was about 5 lbs (2.3 kg). In the flights of 1940, the additional weight of the temperature element was compensated by reducing the length of the rubber shock cord. In all the flights latex balloons, model 700 (weight 700 to 740 g), made by the Dewey and Almy Chemical Co., were used. These balloons were joined in tandem by means of especially strong long-staple cotton twine (regular twist, mainsail stitching twine; 12S/14 ply; weight 130 ft/oz) having a minimum breaking strength of 30 lbs (14 kg). To provide a long period of pendular swing, the distance between the box containing the ultraviolet meter and the point of attachment of the lowest balloon was 30 to 32 m. This suspension cord consisted of three strands of twine having a minimum breaking strength of 90 lbs, which was needed when launching the apparatus in a strong wind.

The distance between the points of attachment of the individual balloons was 8 m; the sections being graded from three strands for the lower three balloons to one strand each for the two topmost balloons. Each balloon was held to the main lifting cord by a 5-m length of this twine.

In some of the flights provision was made for releasing the balloons after bursting, as described by Johnson [13]. However, the device usually became entangled with the main lifting cord and failed to drop off. Moreover, in all recoveries of apparatus, the main part of the balloon was blown into long narrow strips (0.5 to 2.5 cm wide; 0.5 to 1.5 m long) that sometimes became entangled in the lifting cord, and only the heavy tube at the end (weighing about 1.5 to 2 oz; 40 to 60 g) remained. The additional "dead" weight on the bursting of one balloon, therefore, had but little effect upon the flight of the apparatus, which is determined by the lifting power of the remaining balloons. Hence, this refinement in making flights was not continued.

All of the ascents were made at Beltsville, Md. In 1939 three flights were made, and there were three recoveries. In the first flight (June 22), using Cd-38303, two balloons were lost soon after the start; therefore it was a four-balloon ascent, to 78,000 ft, at the rate of about 630 ft./min. The descent by two balloons began at 1:00 p. m. The radio sonde was seen falling into Chesapeake Bay, near Baltimore, from which it was recovered about 4 p. m.

In the second flight (June 26, to a height of 78,000 ft), the ascent was at the rate of 1,325 ft/min with six balloons; then 875 ft/min with five balloons; then 550 ft/min with four balloons—descent on three balloons.

In the third flight (July 1, 1939) using seven balloons, each having a free lift of 1 lb 14 oz, a height of 87,000 ft (perhaps 89,000 ft) was attained—one-balloon descent. The rate of ascent was 920 ft/min to an elevation of 70,000 ft, then 670 ft/min. The ascent started at 10:40 a. m. and the descent at 12:43 p. m. The apparatus was recovered at 2:10 p. m. near Middletown, Del., a distance of about 70 miles from the starting point. In the two previous flights, the

apparatus was recovered in nearby Maryland and Virginia, respectively.

In 1940, four ultraviolet radio sonde flights were made, only three of which were successful. In each flight six balloons were used, each balloon being inflated to lift 1 lb 14 oz. In the fourth flight (June 27), one balloon became detached during the launching; hence it was a five-balloon ascension.

The first flight (June 17) was unsuccessful because of local radio interference, and the radio transmitter may have been injured in launching the apparatus. The apparatus was recovered the following morning by a coast guard on Cape Henlopen Beach, N. J.—three balloons intact, box crushed. Apparently the partly inflated balloons had dragged the box along the shore, breaking all the apparatus.

In the second flight (June 20), using Cd-38303, the ascent began at 10:59 a. m. in a strong breeze. The maximum elevation (83,000 ft) was attained at 12:47 p. m. The descent was followed until 4:30 p. m., when the sun was too low to shine upon the ultraviolet meter but seemed to heat it (see fig. 1). The flight record is interesting in that the measurements of ultraviolet intensities at the ceiling give a more sharply defined indication of the time the highest elevation was attained than the barograph record. As mentioned on a preceding page, during the descent the apparatus floated for some hours at an elevation of about 70,000 ft, and was blown out to sea beyond Boston, where it fell into the ocean 24 hours later. Whether the box of apparatus was broken on the shoals, or by grappling hooks in making the recovery, was not learned. As already noted, in the flights of 1939 this photoelectric cell fell into Chesapeake Bay.

In the third flight (June 22) the ascent started at 10:30 a. m. The maximum elevation (87,300 ft) was attained at 12:25 p. m. In this flight, as in the preceding one, the rapid change in ultraviolet intensities with elevation gave a more sharply defined indication of the time the maximum elevation was attained than the barographic record. In this flight and in the following one there was no recovery of apparatus. Presumably this was owing to (upper air) weather conditions which were unusual, in that there were but few clear days all summer; and presumably, also, owing to the fact that the balloons did not descend promptly the same day (or in the daytime) when the descent could be seen. Nevertheless, as a whole, the score of recoveries of apparatus (14 recoveries out of 19 flights; 74 percent in 4 years) seems remarkably high, considering the fact that the flights were made so close to large bodies of water, marshy areas, and infrequently visited lowland covered with trees and underbrush. Of the 14 recoveries, three were from the water and two on the beach, bordering large bodies of water.

III. EVALUATION OF THE ULTRAVIOLET INTENSITY MEASUREMENTS

For details regarding the methods of evaluation of these measurements, reference is made to preceding papers [1, 5, 6]. It will, therefore, suffice to recall that the evaluation of solar ultraviolet intensities in the stratosphere requires a knowledge of: (a) the spectral-response curves of the photoelectric cell; (b) the spectral ultraviolet transmission of the filters; (c) the spectral-absorption coefficients of ozone

and Rayleigh scattering; (d) the solar-energy curve outside the atmosphere; and (e) the change in transmission of the integrated ultraviolet solar radiation, resulting from a change in temperature of the photoelectric cell and filters used in the ultraviolet meter. The latter enters as a minor correction in the calculations.

1. EFFECT OF TEMPERATURE ON THE ULTRAVIOLET MEASUREMENTS

The effect of temperature on the ultraviolet solar-radiation measurements was determined by enclosing an ultraviolet meter in a heat-insulated, airtight box that was cooled with solid CO_2 ("dry ice"). The source of ultraviolet was a tungsten Mazda *CX* lamp, operated under conditions to give filter transmissions of integrated ultraviolet radiation closely similar to those obtained in measuring ultraviolet solar radiation in the stratosphere (see fig. 2 of reference [11]).

Starting with the apparatus at room temperature (25° C), filter transmissions were made at decreasing temperatures, down to −40° C. By this procedure the apparatus was always warmer than the refrigerant, and no moisture condensed on the windows and the filters.

In the temperature range of +25° to −35° C(Δ=60° C) the filter transmissions increased by 2.1 to 2.8 percent (average 2.5 percent for all photocells) of the value observed at room temperature. For example, using photoelectric cell Cd-38303, the transmissions of the *LG*-filter increased from 21.0 to 23.8 percent and the *LS*-filter increased from 42.2 to 45.0 percent (Δ=2.8 percent).

This is equivalent to a decrease in spectral sensitivity of the photoelectric cell in the short wave lengths (or increase in sensitivity in the long wavelengths with decrease in temperature), or an increase in spectral transmission of the filters in the short wavelengths as observed by Arrol [18]. The combined effect of these two elements is the opposite of the change in spectral quality of the ultraviolet (increase in intensity in the short wavelengths, producing a decrease in percentage transmission of the filters) observed with increase in elevation above the earth's surface. This correction is small, however; and, because of the rise in temperature (see fig. 1), is practically negligible at the highest elevations. This test for change in spectral response is extremely sensitive and easily observed. The change in total response (intrinsic sensitivity) was too small to separate from the errors of measurement of the total intensity.

In figure 1 are depicted the measurements of the ambient temperature of the photoelectric cell and the filters, observed on two flights into the stratosphere. It is interesting to note that the lowest temperatures (−15° C) within the box occurred at an elevation of 15 km, which was the starting point in the calculations of our earliest measurements [5, 6] and is the beginning of an appreciable change in concentration of ozone per kilometer as shown in figure 5.

Incidentally, it is relevant to note that, relative to the large variations in the intensity measurements, the corrections for temperature are small (largest at 12 to 16 km), so that by a fortuitous combination of data on spectral response of the photocells and transmissions of the filters, used in the stratosphere measurements of 1938, the graph (fig. 14 of RP1207 [1]) showing the distribution of ozone at various elevations above the earth's surface is in good agreement with subsequent observations (see fig. 5).

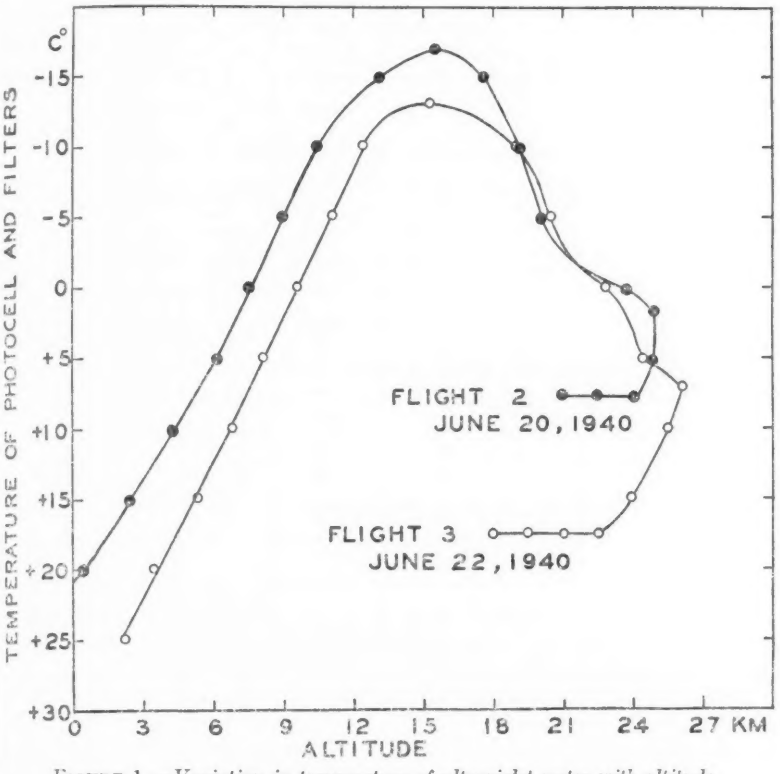


FIGURE 1.—*Variation in temperature of ultraviolet meter with altitude.*

2. THE EXTRATERRESTRIAL ULTRAVIOLET SOLAR SPECTRAL-ENERGY CURVE

An important factor that enters into the calculations of the ozone distribution above the apparatus is the shape of the curve depicting the solar spectral-energy curve of wavelengths shorter than about 3200 A, outside the earth's atmosphere. The solar spectrograms and spectrum photographic density measurements of Fabry and Buisson [12] show large Fraunhofer bands in the region of 2900, 2940, 3025, and 3120 A, with two bands of high intensity in the region of 2985 and 3055 A, respectively. The early measurements of Pettit, using integrated sunlight, indicated a high emission in the region of 2980 A; but this was not observed on the center of the solar disk [4]. All our measurements by means of photoelectric cells and filters [14, 11] indicate a high emission in the region of 2980 A. A blueprint of unpublished observations by Brian O'Brien (given us in 1930) shows a high emission in the region of 2950 A and a low emission at 3030 A.

In view of these uncertainties, and in order to obtain calculated filter transmissions that are in agreement with the observed filter transmissions of the integrated ultraviolet incident upon the photoelectric cell, we adopted a spectral-energy distribution that is high in the region of 2980 A (see table 1) and low in the region of 3050 A—

omitting the narrow emission at 3055 A shown in the spectrograms of Fabry and Buisson, in order to simplify the calculations. Incidentally, it is to be noted that the various observers are in fair agreement in the evaluation of the energy distribution in the spectral region of 3100 to 3300 A.

TABLE 1.—*Solar energy outside the earth's atmosphere*

[From all data available: Pettit (integrated and center of disk, 1931-36); Fabry and Buisson (1922); O'Brien (unpublished 1930, blueprint); Coblentz and Stair (1935); Stair and Hand (1939)]

Wavelengths	Energy	Wavelengths	Energy	Wavelengths	Energy	Wavelengths	Energy
A		*A*		*A*		*A*	
2800	0.5	3000	75	3200	90	3400	97.0
20	1.3	20	70	20	93	20	96
40	3.0	40	54	40	95	40	96
60	5.3	60	54	60	96	60	96
80	9.0	80	56	80	97	80	96.5
2900	15.0	3100	60	3300	97.5	3500	97
20	30	20	66	20	98	20	98
40	60	40	72	40	98	40	100
60	73	60	76	60	98	60	103
80	75	80	83	80	97.5	80	106
						3600	110
						3700	112

In calculating the distribution of ozone in the stratosphere, use is made of the solar spectral-energy data given in table 1, and the spectral-absorption coefficients of ozone [16] described in previous papers [1, 6].

While numerous determinations of the spectral-absorption coefficients of ozone in the atmosphere are in close agreement with similar measurements on ozone prepared in the laboratory [15], it is to be noted that recent investigations by Adel and Lampland [8] seem to indicate the presence of oxides of nitrogen (N_2O and N_2O_5) in the upper atmosphere [9], produced by the photochemical action of ultraviolet solar radiation. Adel [9] points out that, if the coincidence of the bands of selective absorption (at 7.77 and 8.57 μ, respectively) in the atmosphere and of laboratory preparations of the above-mentioned oxides of nitrogen is not fortuitous, there would appear to be a layer of N_2O, comparable to the known layer of ozone, in the stratosphere. Such a deduction has wide implications. The writers do not recall data on the ultraviolet spectral absorption of N_2O. The measurements of Jones and Wulf [10] show no bands of selective absorption of N_2O_5 in the ultraviolet; but an increase in absorption with decrease in wavelength is recorded, the spectral-absorption coefficients being similar in magnitude to those of ozone. It is therefore an interesting scientific question how much of the observed ultraviolet spectral absorption is caused by each of these gases (O_3, N_2O, N_2O_5); though for our purposes it seems relatively unimportant, since use is made of the directly observed spectral-absorption coefficients that appear to be representative of all of the gases that form the absorbing layer in the stratosphere.

To recapitulate, using sunlight, also starlight [15], as a source, measurements show narrow bands of selective absorption superposed upon a wide band of general absorption in the terrestrial atmosphere, that increases with decrease in wavelength in the extreme ultraviolet

of the solar-energy spectrum. Earlier measurements of these atmospheric ultraviolet-absorption coefficients have been correlated with definite amounts (n. t. p.) of laboratory preparations of ozone. Recent infrared solar spectral-energy measurements show (atmospheric) absorption bands identifiable with laboratory preparations of certain oxides of nitrogen. If this is not a fortuitous, approximate, coincidence (because of the small dispersion usable) of Fraunhofer

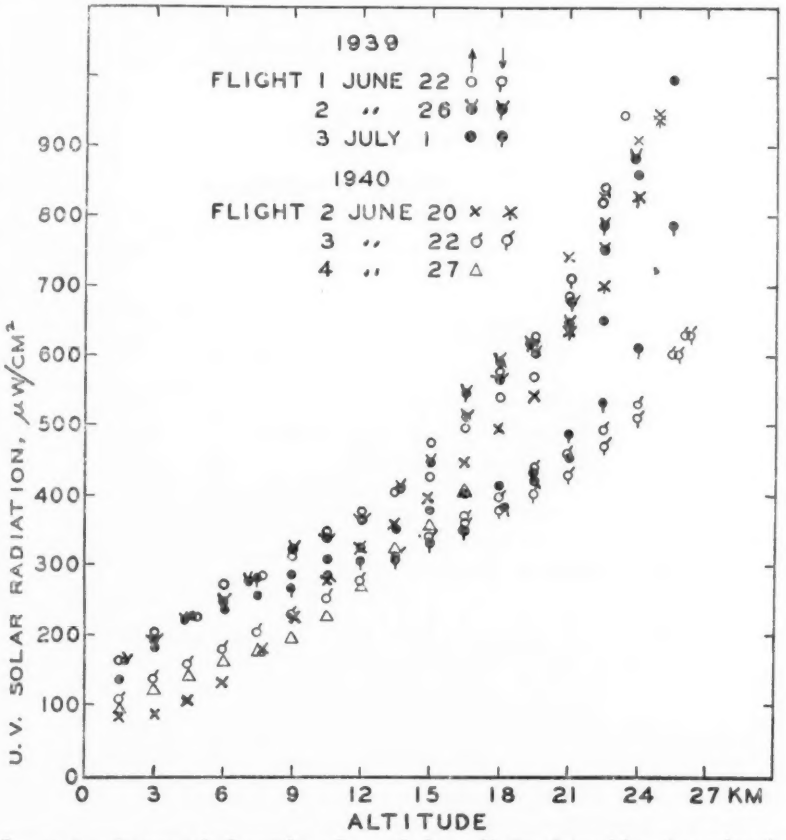


FIGURE 2.—*Integrated ultraviolet solar and sky radiation intensities of wavelengths shorter than 3132 A at various altitudes above the earth's surface.*

lines, the presence of oxides of nitrogen in the upper atmosphere is to be inferred, and is to be expected on the basis of ultraviolet photochemical action. However, these deductions and inferences do not modify the herein-described measurements and conclusions, which pertain to the position, and not the composition of the absorbing layer, in the stratosphere.

3. EXPOSITION OF DATA

The ultraviolet intensities are received at the ground station in the form of a graphical record of audio frequencies emitted by the photoelectrically controlled radio transmitter, during ascent and descent of

the apparatus, as shown in figure 8 of the preceding paper [1]. From such records of ascents in 1939 and 1940, data are obtained showing the variation in ultraviolet solar intensity with altitude, as depicted in figures 9 and 10 of the preceding paper [1].

In figure 2 are depicted the integrated ultraviolet intensities of ultraviolet solar radiation of wavelengths shorter than 3132 A, at various elevations above sea level, observed in 1939 and in 1940. At the lower elevations the wide spread in the intensity measurements is to be attributed to clouds, haze, and dust on different days. At the

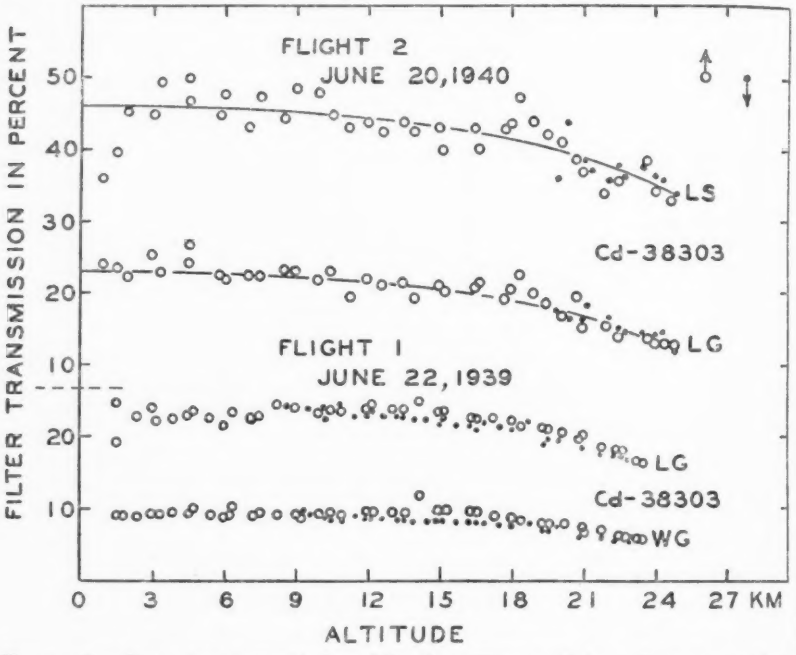


FIGURE 3.—*Percentage transmission of the glass filters used in various stratosphere flights.*

highest elevations the ultraviolet intensities observed in 1939 and in 1940 are in good agreement with the measurements of 1938, indicating an intensity of about $900\mu W/cm^2$ of ultraviolet radiation of wavelengths shorter than 3132 A, at an altitude of about 27 km [1].

From data similar to those depicted in figures 9 and 10 [1], the filter transmissions are calculated by two methods as outlined on page 599 and depicted in figure 11 of the preceding publication [1]. Some of these filter transmissions, observed at different elevations of the ultraviolet meter, using photoelectric cell Cd–38303, are plotted in figure 3. The arrows in the upper right-hand corner of this illustration indicate the ascent and descent of the apparatus. In order to increase the accuracy of the audio-frequency readings, the glass filters used in the flights of 1940 had higher integrated ultraviolet transmissions than those used in 1938 and in 1939. Using the same photoelectric cell (Cd–38303), it may be noted that the individual observations of 1940 suffered greater variations than those obtained with this apparatus

during the two preceding years. In the case of this particular photoelectric cell this variation may have been caused partly, perhaps, by variation in contact resistance of the electrode with the photosensitive surface, which seemed to vary somewhat after recovery of the apparatus, in 1940.

These filter transmissions at different elevations have been corrected for temperature, see figure 1. From these filter-transmission curves [16] the amount of ozone traversed at a given height above the earth's

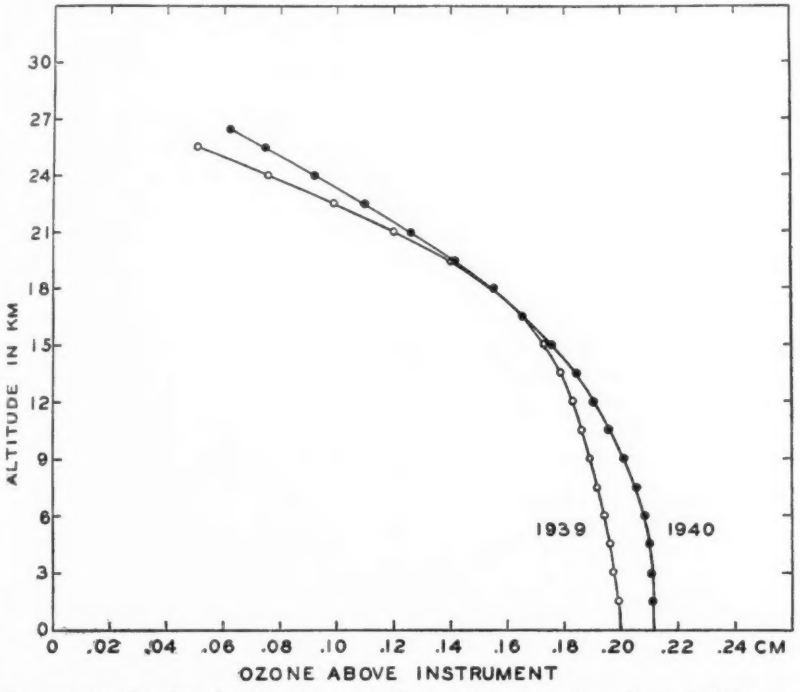


FIGURE 4.—*Graphs showing the amount of ozone above the ultraviolet-intensity meter, at various heights above the earth's surface, observed in the flights of 1939 and 1940.*

surface is then calculated, as outlined on page 602 of the previous paper [1].

In figure 4, curves are depicted showing the amount of ozone above the ultraviolet-intensity meter, at various heights above the earth's surface. The graph for 1939 is the average of three flights, and that of 1940 is the average of two flights (see fig. 2 for the dates of the flights). Owing to the difficulties in making such observations, the difference between the single curves and the average of the two curves, at any elevation, is not sufficiently marked to be considered real.

Following the procedure previously outlined (p. 604 [1]), from figure 4 is derived the amount of ozone per kilometer (O_3/km) at different elevations above the earth's surface.

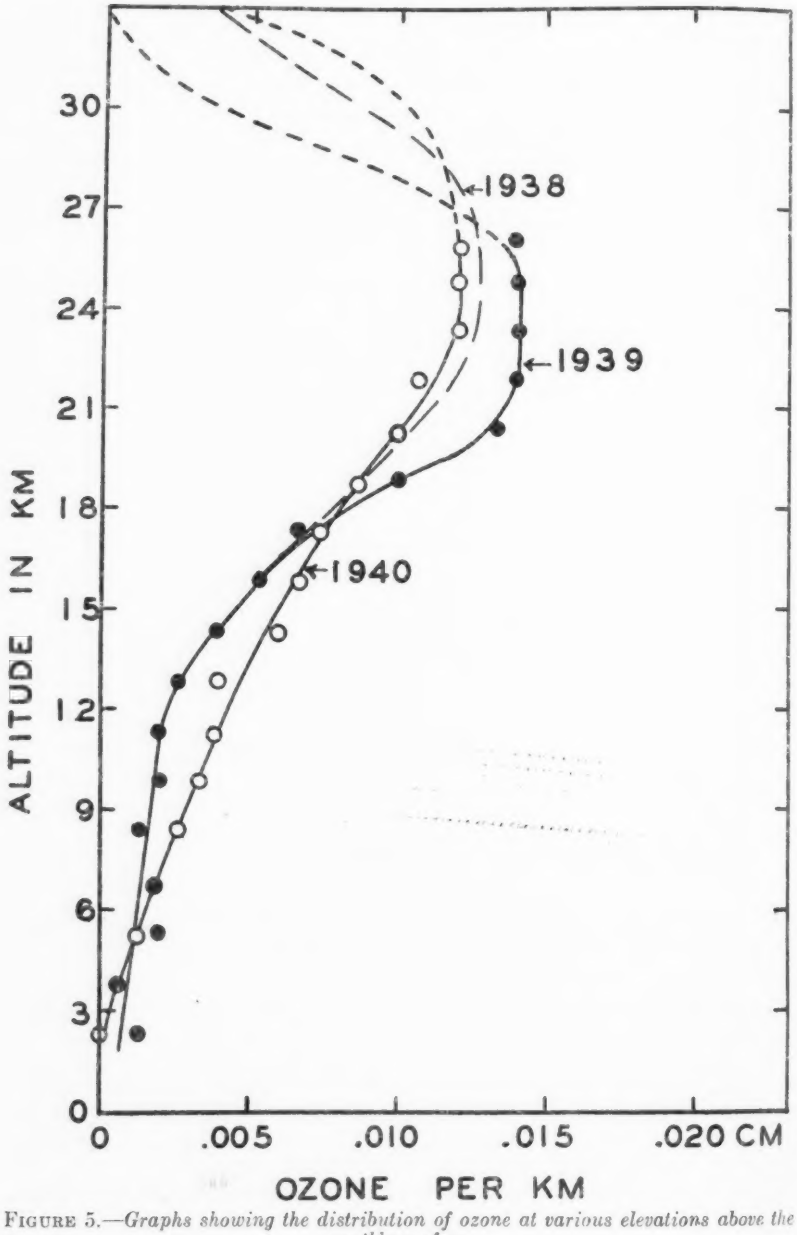


FIGURE 5.—*Graphs showing the distribution of ozone at various elevations above the earth's surface.*

As shown in figure 5, the relatively high concentration of ozone may perhaps have extended to lower elevations (12 km) in 1940 than in 1939. In this illustration is depicted also the curve obtained in 1938, after correcting the filter transmission for changes in temperature of the apparatus during the flights. The explorations of all three seasons are in good agreement in placing the maximum concentration of the ozone at a height of 23 to 25 km. The dotted part of these curves shows the distribution of ozone required to account for the 30 to 35 percent of ozone that still remained above the apparatus when it was at an elevation of about 27 km above sea level.

IV. REFERENCES AND NOTES

[1] W. W. Coblentz and R. Stair, *Distribution of ozone in the stratosphere*, J. Research NBS **22,** 573 (1939) RP1207.
A complete report of the stratosphere flights, and of ultraviolet measurements at Flagstaff, Ariz. (elevation 7,300 ft) in 1938.

[2] D. N. Craig, *Electrolytic resistors for direct-current applications in measuring temperature*, J. Research NBS **21,** 225 (1938) RP1126.

[3] H. Diamond, W. S. Hinman, F. W. Dunmore and E. G. Lapham, *An improved radio sonde and its performance*, J. Research NBS **25,** 327 (1940) RP1329; also RP1082 (1938).
See figure 5 for electrolytic temperature-measuring element, samples of which were supplied us by Julien P. Friez & Sons, Inc., Baltimore, Md.

[4] E. Pettit, *Spectral energy curve of the sun in the ultraviolet*, Astrophys. J. **91,** 159 (1940).

[5] W. W. Coblentz and R. Stair, *A radiometric method of measuring ultraviolet solar-radiation intensities in the stratosphere*, Bul. Am. Meteorolog. Soc. **18,** 345 (1937); Radiologica **1,** 12 (1937). *See also* J. Research NBS **17,** 1 (1936) RP899.
A brief description of the methods used and a preliminary report of the results obtained in the stratosphere flights of 1937.

[6] R. Stair and W. W. Coblentz, *Radiometric measurements of ultraviolet solar intensities in the stratosphere*, J. Research NBS **20,** 185 (1938) RP1075.
Complete report of the results of the stratosphere flights of 1937.

[7] R. Stair, *A precision radio instrument for transmitting measurements of ultraviolet intensities from unmanned balloons to a ground station*, J. Research NBS **22,** 295 (1939) RP1181.
A description of improvements in the radio sonde previously used [5, 6]. By combining a balanced direct-current amplifier [RP647] with a relaxation oscillator and short-wave transmitter similar to that developed by Diamond and his collaborators (RP1082 [3]) measurements of ultraviolet intensities are obtained to the highest altitudes attained by the radio sonde [1].

[8] A. Adel and C. O. Lampland, *A new band in the absorption spectrum of the earth's atmosphere*, Astrophys. J. **87,** 198 (1938).

[9] A. Adel, *Note on the atmospheric oxides of nitrogen*, Astrophys. J. **90,** 627 (1939).

[10] E. J. Jones and O. R. Wulf, *The absorption coefficient of nitrogen pentoxide in the ultraviolet, and the visible absorption spectrum of* NO_3, J. Chem. Physics **5,** 873 (1937).

[11] R. Stair and I. F. Hand, *Methods and results of ozone measurements over Mount Evans, Colorado*, Mo. Weather Rev. **67,** 331 (1939).

[12] C. Fabry and H. Buisson, *Distribution of energy in the ultraviolet of the solar spectrum*, Compt. Rend. **175,** 156 (1922); Astrophys. J. **54,** 297 (1921).
Spectrogram reproduced in J. Research NBS **15,** 123 (1935) RP816.

[13] T. H. Johnson and J. G. Barry, Phys. Rev. **56,** 219 (1939).

[14] W. W. Coblentz and R. Stair, *Distribution of energy in the extreme ultraviolet of the solar spectrum*, J. Research NBS **17,** 1 (1936) RP899.

[15] F. W. P. Goetz, *Die vertikale Verteilung des atmosphärischen Ozons*. Ergebn. Kosmischen Physik III, 253–326, Leipzig, 1938.
An up-to-date summary of the researches on atmospheric ozone including 21 references to his own work, published since 1931. His survey of previous investigations on atmospheric ozone is published in the "Ergebnisse" **I,** 180–235 (Leipzig, 1931).

[16] C. Fabry and H. Buisson, *Data on ozone absorption*, Compt. rend. **192,** 457 (1931).
Ny Tsi-Ze et Choong Shin-Piaw, Compt. rend. **195,** 309 (1932); **196,** 916 (1933).
F. W. Paul Goetz, Ergebn. Kosmischen Physik **3,** 253 (1938).
[17] W. W. Coblentz and R. Stair, *Evaluation of ultraviolet solar radiation of short wave lengths*, J. Research NBS **16,** 315 (1936) RP877.
[18] W. J. Arrol, *Change of ultraviolet transparency of glass with temperature*, Nature **145,** 861 (1940).
In the range of +10° C to −50° C the absorption coefficient was observed to increase linearly with increase in temperature.

WASHINGTON, November 1, 1940.